DISSERTATION | DOCTORAL THESIS

Titel | Title

Towards a Comparative Approach to Empirical Aesthetics: Complexity and Emotion in Environmental Scenes, Representational Paintings and Music

verfasst von | submitted by

Mag.phil. Mag.phil. Manuela Marin MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 298
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Psychologie
Betreut von | Supervisor:	Univ.-Prof. Dipl.-Psych. Dr. Helmut Leder

This is in memory of Léa.

Cézanne’s Ports

Allen Ginsberg

In the foreground we see time and life

swept in a race

toward the left hand side of the picture

where shore meets shore.

But that meeting place

isn’t represented;

it doesn’t occur on the canvas.

For the other side of the bay

is Heaven and Eternity,

with a bleak white haze over its mountains.

And the immense water of L’Estaque is a go-between

for minute rowboats.

Acknowledgements

Embarking on a PhD journey in psychology is more than just striving for academic achievement. It is about professional and personal growth, and it is about acknowledging that science is a team effort. I am indebted to numerous people who have accompanied me on my journey and supported and encouraged me over the years. First and foremost, I would like to express my sincere gratitude to Prof. Helmut Leder, my supervisor who always had profound belief in my abilities and whose guidance and patience was essential for the completion of my doctoral thesis. I also collected unforgettable memories as a member of his research group in Vienna between 2011 and 2015. I would like to thank my former colleagues for many fruitful discussions of our research projects and their constructive feedback. I would like to especially thank Dr. Andreas Gartus, whose technical assistance and humour was greatly appreciated over the years. I am also grateful to Dr. Martina Jakesch with whom I shared an office and many enjoyable moments. I also had great pleasure of training and working with many students (interns and Master's students) during these productive years, and without their assistance large-scale experimental work in the laboratory would not have been possible. Special thanks to a large number of participants who provided us with the data on which this doctoral thesis is based on. Since the research articles of this doctoral thesis are already published, I would like to thank the (anonymous) reviewers for their critical readings and invaluable insights.

I would also like to thank all those students/technicians/colleagues who helped me along the way but whose names are too numerous to be mentioned here.

At a more personal level, I wish to thank my husband Dr. Bruno Gingras for his unwavering support and relentless insistence to move forward in my career. I am also largely indebted to my family, and especially to my late maternal grandmother, who always supported and nurtured me until 2021. I will always miss her. Finally, I am extremely grateful to a team of Tyrolean doctors who managed to cure me between 2021 and 2022 from a long-hidden disease.

I dedicate this doctoral thesis to my late mother-in-law, Prof. Léa Brakier-Gingras (1940-2021), in gratitude for her encouragement over many years. Léa was an excellent professor, loved by her students, and thus has inspired many generations of academics, including me.

Abstract

The field of empirical aesthetics is nearing its 150th anniversary and is more prolific than ever in terms of its research output. Several competing theoretical models based on cognitive and affective processing have attempted to explain how aesthetic judgements and emotions emerge during aesthetic episodes. Most of these models posit that aesthetic responses are influenced by a combination of the characteristics of the stimulus, as well as person-related and contextual factors, but they differ regarding the explanatory psychological mechanism. In this doctoral thesis, we propose that a comparative approach, involving stimuli of different sensory modalities, is effective and valuable in determining the nature of aesthetic experiences as well as the domain-general and/or domain-specific aspects thereof (Marin, 2015). Here, we systematically studied the experience of affective environmental scenes, representational paintings and music in the context of Berlyne's psychobiological model (1971), which is still widely cited within the field and beyond, but which has also led to conflicting results. The affective content of stimuli, although frequently occurring and relevant in our daily lives, has often been ignored in recent research, together with the study of subjective arousal (Marin, 2022). Thus, we studied stimulus complexity as well as felt pleasantness and arousal. Following a multi-method approach, we collected subjective responses, employed computational methods to analyze our stimuli (Marin & Leder, 2013), and also measured eye movements in a laboratory setting (Marin & Leder, 2022). We examined the dynamics of visual aesthetic experiences by using different presentation durations ranging between 1 and 25 s (Marin & Leder, 2016), and studied individual differences rooted in affective personality traits and art/music interest (Marin & Leder, 2018), in addition to scrutinizing Berlyne's concept of hedonic tone (Marin et al., 2016). The main insights gained from this comprehensive research program were that visual stimuli followed a U-shaped relationship between pleasantness and arousal in a two-dimensional emotion space, whereas musical stimuli were distributed in a circular/elliptical shape. After controlling for familiarity, complexity and arousal were more strongly correlated when viewing artworks and listening to music than when viewing environmental scenes. Berlyne's theory was further supported by demonstrating mediation effects of arousal in the complexity-pleasantness relationship for all stimulus categories. However, there was very little support for the predicted inverted U-shaped relationship between complexity and pleasantness. Further investigations showed that hedonic tone, as described

by Berlyne, is not a uniform concept. Whereas beauty, liking and pleasantness were highly correlated when viewing environmental scenes, these ratings were less correlated when viewing artworks or listening to music, and moreover, they showed different correlation patterns with complexity, which may partly explain earlier discrepant research findings. Furthermore, we computed 13 objective measures of visual complexity (mostly based on Kolmogorov complexity and edge detection) and demonstrated that the strength of the correlations of these measures with subjective complexity depended on the stimulus type. A linear combination of different measures yielded the best results, with RMS contrast measures emerging as the strongest predictors for environmental scenes and phase congruency measures for representational paintings. Additionally, we showed that these types of complexity measures were also applicable to music, with compressed file sizes and event density yielding strong correlations with subjective complexity and arousal. Objective measures of complexity were also able to capture the perceived subdimensions of visual complexity (i.e., the number of elements vs. disorganization), and these correlations were less affected by presentation duration than correlations with general complexity. The multidimensionality of visual complexity for environmental scenes and paintings was shown to be more clearly present for a viewing time of 5 s than 1 s. Our highly-controlled eye movement study revealed that pleasantness was the main factor differentiating between the viewing behaviour of environmental scenes and representational paintings. In general, most of our findings were affected by presentation duration. Regarding individual differences, affect-related personality traits were related to complexity ratings of artistic stimuli and to arousal ratings of all stimulus categories, whereas art interest and musical sophistication were found to play a minor role among females. Our findings were discussed within the context of Berlyne's motivational theory, which was partly supported by our data, and current alternative theories. We also elaborated on the nature and dynamics of aesthetic emotions and aesthetic experiences, and complexity and emotion research more generally. In conclusion, our comparative approach has yielded novel insights that would have been difficult to obtain by focusing on a single stimulus type. Therefore, we encourage researchers to replicate and extend our findings by studying a wider range of artistic and sensory stimuli under a common framework, incorporating the timely study of individual differences and the dynamics of aesthetic experiences, preferably also in a variety of social and cultural contexts.

Zusammenfassung

Das Forschungsfeld der empirischen Ästhetik nähert sich seinem 150-jährigen Bestehen und ist in Bezug auf seinen Forschungsoutput produktiver denn je. Mehrere konkurrierende theoretische Modelle, die auf kognitiver und affektiver Verarbeitung basieren, haben versucht zu erklären, wie ästhetische Urteile und Emotionen während ästhetischer Episoden zustande kommen. Die meisten dieser Modelle gehen davon aus, dass ästhetische Reaktionen durch eine Kombination aus den Merkmalen des Reizes und personen- sowie kontextbezogenen Faktoren beeinflusst werden, aber sie unterscheiden sich hinsichtlich des erklärenden psychologischen Mechanismus. In dieser Dissertation schlagen wir vor, dass ein vergleichender Ansatz, der Stimuli verschiedener sensorischer Modalitäten einbezieht, effektiv und wertvoll ist, um die Natur ästhetischer Erfahrungen sowie deren domänenübergreifende und/oder domänenspezifische Aspekte zu bestimmen (Marin, 2015). Hier haben wir das Erleben von realen, emotionalen Umweltszenen, gegenständlichen Gemälden und Musik systematisch im Kontext des psychobiologischen Modells von Berlyne (1971) untersucht, das in diesem Bereich und darüber hinaus immer noch häufig zitiert wird, aber auch zu widersprüchlichen Ergebnissen geführt hat. Der affektive Gehalt von Reizen, obwohl er häufig vorkommt und in unserem täglichen Leben von immenser Bedeutung ist, wurde in der neueren Forschung oft ignoriert, ebenso wie die Untersuchung der subjektiven Erregung (Marin, 2022). Wir interessierten uns daher insbesondere für die Komplexität der Stimuli sowie für das empfundene Wohlgefallen (angenehm vs. unangenehm) und ihren Anregungsgehalt. Wir verfolgten einen multimethodalen Ansatz, indem wir subjektive Beurteilungen sammelten, computergestützte Methoden zur Analyse unserer Stimuli einsetzten (Marin & Leder, 2013) und auch Augenbewegungen in einer Laborumgebung maßen (Marin & Leder, 2022). Wir untersuchten die Dynamik visuell-ästhetischer Erfahrungen mit unterschiedlichen Präsentationsdauern zwischen 1 und 25 Sekunden (Marin & Leder, 2016), des Weiteren individuelle Unterschiede, die auf affektiven Persönlichkeitsmerkmalen sowie auf Kunst- und Musikinteresse beruhen (Marin & Leder, 2018), und hinterfragten Berlynes Konzept des hedonischen Tons (Marin et al., 2016). Die wichtigsten Erkenntnisse aus diesem umfassenden Forschungsprogramm waren, dass visuelle Stimuli einer U-förmigen Beziehung zwischen Wohlgefallen und Erregung in einem zweidimensionalen Emotionsraum folgten, während musikalische

Stimuli in einer kreisförmigen/elliptischen Form verteilt waren. Wenn für die Vertrautheit kontrolliert wurde, waren Komplexität und Erregung beim Betrachten von Kunstwerken und beim Hören von Musik stärker korreliert als beim Betrachten von realen Umweltszenen. Berlynes Theorie wurde außerdem durch den Nachweis von Mediationseffekten von Erregung in der Beziehung zwischen Komplexität und Wohlgefallen für alle Reizkategorien unterstützt. Die vorhergesagte umgekehrte U-förmige Beziehung zwischen Komplexität und Wohlgefallen wurde jedoch nur in sehr geringem Maße unterstützt. Weitere Untersuchungen zeigten, dass der hedonische Ton, wie er von Berlyne beschrieben wurde, kein einheitliches Konzept ist. Während Schönheit, Gefallen und Wohlgefallen beim Betrachten von Umweltszenen hoch korreliert waren, waren diese Bewertungen beim Betrachten von Kunstwerken oder beim Hören von Musik weniger stark korreliert, und außerdem zeigten sie unterschiedliche Korrelationsmuster mit der Komplexität, was frühere diskrepante Forschungsergebnisse teilweise erklären könnte. Darüber hinaus berechneten wir 13 objektive Maße der visuellen Komplexität (meist basierend auf Kolmogorov-Komplexität und Kantendetektion) und zeigten, dass die Stärke der Korrelationen dieser Maße mit der subjektiven Komplexität von der Art des Stimulus abhing. Eine lineare Kombination verschiedener Maße lieferte die besten Ergebnisse, wobei sich RMS-Kontrastmaße als die stärksten Prädiktoren für Umweltsszenen und Phasenkongruenzmaße für gegenständliche Gemälde herausstellten. Darüber hinaus zeigten wir, dass diese Arten von Komplexitätsmaßen auch auf Musik anwendbar sind, wobei komprimierte Dateigrößen und Ereignisdichte starke Korrelationen mit subjektiver Komplexität und Erregung ergaben. Objektive Komplexitätsmaße waren auch in der Lage, die wahrgenommenen Unterdimensionen der visuellen Komplexität zu erfassen (d.h. die Anzahl der Elemente und die Desorganisation), und diese Korrelationen wurden von der Präsentationsdauer weniger beeinflusst als die Korrelationen mit allgemeinen Komplexitätsbewertungen. Die Mehrdimensionalität der visuellen Komplexität von Umweltszenen und Gemälden zeigte sich bei einer Betrachtungsdauer von 5 s deutlicher als bei 1 s. Unsere hochgradig kontrollierte Augenbewegungsstudie ergab, dass das Wohlgefallen der Hauptfaktor war, der zwischen dem Betrachtungsverhalten von Umweltszenen und gegenständlichen Gemälden unterschied. Im Allgemeinen wurden die meisten unserer Ergebnisse von der Präsentationsdauer beeinflusst. Was die individuellen Unterschiede betrifft, so

standen bei allen Stimuluskategorien affektbezogene Persönlichkeitsmerkmale im Zusammenhang mit der Bewertung der Komplexität künstlerischer Stimuli und mit der Bewertung der Erregung, während Kunstinteresse und musikalische Erfahrenheit bei Frauen eine untergeordnete Rolle spielten. Unsere Ergebnisse wurden im Zusammenhang mit der Motivationstheorie von Berlyne, die durch unsere Daten teilweise gestützt wurde, und aktuellen alternativen Theorien diskutiert. Außerdem haben wir uns mit der Natur und der Dynamik ästhetischer Emotionen und ästhetischer Erfahrungen sowie mit der Komplexitäts- und Emotionsforschung im Allgemeinen befasst. Zusammenfassend lässt sich sagen, dass unser vergleichender Ansatz zu neuen Erkenntnissen geführt hat, die durch die Konzentration auf einen einzigen Stimulustyp nur schwer zu gewinnen gewesen wären. Daher ermutigen wir Forscher:innen, unsere Ergebnisse zu replizieren und zu erweitern, indem sie ein breiteres Spektrum künstlerischer und sensorischer Stimuli in einem gemeinsamen Rahmen untersuchen, der die aktuelle Untersuchung individueller Unterschiede und der Dynamik ästhetischer Erfahrungen einschließt, vorzugsweise auch in einer Vielzahl sozialer und kultureller Kontexte.

Publications included in this doctoral thesis

Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLOS One*, *8*(8), Article e72412. https://doi.org/10.1371/journal.pone.0072412

Marin, M. M. (2015). Crossing boundaries: Toward a general model of neuroaesthetics. *Frontiers in Human Neuroscience*, *9*, Article 443. https://doi.org/10.3389/fnhum.2015.00443

Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica*, *163*, 38-58. https://doi.org/10.1016/j.actpsy.2015.10.002

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience, 10*, Article 536. https://doi.org/10.3389/fnhum.2016.00536

Marin, M. M., & Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Personality and Individual Differences*, *125*, 80-90. https://doi.org/10.1016/j.paid.2017.12.027

Marin, M. M., & Leder, H. (2022). Gaze patterns reveal aesthetic distance while viewing art. *Annals of the New York Academy of Sciences, 1514*(1), 155-165. https://doi.org/10.1111/nyas.14792

Marin, M. M. (2022). The role of collative variables in aesthetic experiences. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (pp. 385-429). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.20

Conference contributions related to this doctoral thesis

Marin, M. M., & Leder, H. (2012). Zusammenhänge zwischen visueller Komplexität und Wohlgefallen bei der Wahrnehmung realer Umweltszenen und repräsentativer Kunst. Poster presented at the *10. Tagung der Österreichischen Gesellschaft für Psychologie*, Graz (Austria).

Marin, M. M., & Leder, H. (2012). Relating subjective to objective measures of complexity in affective environmental scenes and representative paintings. Talk given at the *European Conference on Visual Perception* (ECVP 2012), Alghero (Italy).

Marin, M. M., & Leder, H. (2013). Effects of arousal, pleasantness and complexity on eye movements when viewing natural scenes and paintings. Talk given at the *55th Conference of Experimental Psychologists* (TeaP 2013), Vienna (Austria).

Marin, M. M., & Leder, H. (2013). Relating subjective and objective measures of musical complexity in different musical styles. Talk given at the *Biennial Meeting of the Society for Music Perception and Cognition* (SMPC 2013), Toronto (Canada).

Marin, M. M., Lampatz, A., & Leder, H. (2013). Berlyne revisited: Evidence for the multidimensional nature of hedonic value in the perception of visual arts. Talk given at the *18th Meeting of the European Society for Cognitive Psychology* (ESCoP 2013), Budapest (Hungary).

Marin, M. M., & Leder, H. (2015). The emergence of the multidimensionality of visual complexity. Poster presented at the *Third Visual Science of Art Conference* (VSAC 2015), Liverpool (UK).

Marin, M. M., Wandl, M., & Leder, H. (2015). Evidence for the multidimensional nature of hedonic value in the perception of Romantic piano music. Talk given at the *Fourth International Conference on Music and Emotion*, Geneva (Switzerland). Talk also given at the *12. Tagung der Österreichischen Gesellschaft für Psychologie*, Innsbruck (Austria).

Marin, M. M., & Leder, H. (2016). Effects of emotion, complexity and presentation duration on eye movements when viewing environmental scenes and paintings. Talk given at the *Fourth Visual Science of Art Conference* (VSAC 2016), Barcelona (Spain).

Marin, M. M., & Leder, H. (2017). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. Talk given at the *Fifth Visual Science of Art Conference* (VSAC 2017), Berlin (Germany).

Table of contents

Dedication
Cézanne's Ports
Acknowledgements
Abstract
Zusammenfassung
Publications included in this doctoral thesis
Conference contributions related to this doctoral thesis

1. Introduction 21

1.1 Theoretical context and research approach 21

1.2 Research questions and methods 36

2. Publications 40

2.1 Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music (Marin & Leder, 2013) 40

2.1.1 Introduction 40

2.1.2 Authorship contribution statement 41

2.1.3 Re-print of the published research article 41

2.1.4 Summary and outlook 77

2.2 Crossing boundaries: Toward a general model of neuroaesthetics (Marin, 2015) 85

2.2.1 Introduction and summary 85

2.2.2 Authorship contribution statement 85

2.2.3 Re-print of the published theoretical article 85

2.3 Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings (Marin & Leder, 2016) 91

2.3.1 Introduction 91

2.3.2 Authorship contribution statement 92

2.3.3 Re-print of the published research article 92

2.3.4 Summary and outlook 114

2.4 Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music (Marin, Lampatz, Wandl, & Leder, 2016) 120

2.4.1 Introduction 120

2.4.2 Authorship contribution statement 121
2.4.3 Re-print of the published research article 121
2.4.4 Summary and outlook 142

2.5 Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures (Marin & Leder, 2018) 146
2.5.1 Introduction 146
2.5.2 Authorship contribution statement 146
2.5.3 Re-print of the published research article 146
2.5.4 Summary and outlook 158

2.6 Gaze patterns reveal aesthetic distance while viewing art (Marin & Leder, 2022) 162
2.6.1 Introduction 162
2.6.2 Authorship contribution statement 162
2.6.3 Re-print of the published research article 162
2.6.4 Summary and outlook 174

2.7 The role of collative variables in aesthetic experiences (Marin, 2022) 178
2.7.1 Introduction and summary 178
2.7.2 Authorship contribution statement 179
2.7.3 Re-print of the published book chapter 179

3. Discussion 225
3.1 Summary of main results 225
3.2 Berlyne revisited: Arousal and complexity still rule 234
3.3 Emotion across sensory and artistic domains 240
3.4 Complexity across sensory and artistic domains 246
3.5 Nature of aesthetic experiences 251
3.6 Conclusions 255

4. References 260

1. Introduction

1.1 Theoretical context and research approach

Empirical aesthetics is the second oldest research field in psychology (after psychophysics), dating back to Gustav Theodor Fechner (1801-1887), who was one of the first experimental psychologists (Jacobsen, 2006). He set the stage for many researchers to follow his interest in determining underlying factors of aesthetic pleasure. Fechner's work on empirical aesthetics, which he considered as part of general hedonics, greatly inspired Daniel E. Berlyne (1924-1976), a motivational psychologist with a strong interest in neurobiology. Berlyne initiated a revival of empirical aesthetics in the second half of the 20th century – the so-called *New Experimental Aesthetics* (Cupchik, 2022; Nadal & Urena, 2022). His arousal-based theory of hedonic response (Berlyne, 1971) has had a long-lasting influence on contemporary researchers in the field, many of whom, for instance, have tried to replicate his famous inverted U-shaped curve describing the relationship between a *collative variable* (e.g., complexity, novelty and ambiguity) and a measure of *hedonic tone* (e.g., liking, beauty and pleasingness). Although Berlyne's focus on structural stimulus properties is still debated (e.g., Martindale & Moore, 1989; Reber et al., 2004; Silvia, 2005), his theory is widely cited by researchers working in different fields, ranging from basic research involving music, visual arts and beyond, to more applied fields, such as website design and consumer behaviour (for a review, see Marin, 2022).

Berlyne (1960, 1971) extended his motivational theory of curiosity and arousal to the study of aesthetic responses. In a nutshell, describing his *arousal-boost mechanism*, he argued that the pleasure experienced when being exposed to an aesthetic object, such as a painting, depends on the degree of arousal which is determined by three types of variables (i.e., *collative*, *psychophysical* and *ecological*). Collative variables play a particular role in his theory, and the arousal induced by these variables, due to seeking out stimulation by exploring the object, activates a primary reward system and a primary aversion system. These systems, Berlyne argued, produce a combined effect of arousal on hedonic tone that follows an inverted U-shaped curve between the respective variables. Specifically, Berlyne predicted a positive linear relationship between a collative variable and arousal, and an inverted U-shaped curve between a collative variable and a measure of *hedonic tone* (for more details, see Marin, 2022).

In the last decades, however, researchers have provided well-cited alternative theories, such as the prototype theory of aesthetic preference (Martindale, 1984), the processing fluency theory of aesthetic pleasure (Reber et al., 2004), the information-processing stage model of aesthetic processing (Leder et al., 2004), as well as appraisal theories of emotional responses to art (Silvia, 2005), to name just a few. These theories are deeply grounded in the study of perceptual and cognitive processes and describe how they are linked to aesthetic judgements and/or emotional responses. Nevertheless, this doctoral thesis is embedded within the framework of Berlyne's theory because of its recurrent appearance in the research literature during the last decades (Marin, 2022). Although new theories on aesthetic processing are emerging (Brielmann & Dayan, 2022; Graf & Landwehr, 2015; Redies, 2015; Specker et al., 2021), Berlyne's theory has never been completely disproven or abandoned (see, e.g., Chmiel & Schubert, 2018; Nadal et al., 2010, but see Makin, 2017, for criticism), which leaves room for further investigations. Importantly, alternative theories, such as processing fluency and its relation to aesthetic liking (Reber et al., 2004) have also been prone to criticism (Graf & Landwehr, 2015). In short, empirical aesthetics currently encompasses a panoply of competing theories, which makes it an attractive research field.

Discrepant research findings in the context of Berlyne's theory may partly be rooted in a lack of clear definitions and an incomplete understanding of relevant concepts (see, e.g., Van Geert & Wagemans, 2020). For example, Nadal et al. (2010) demonstrated that the multidimensionality of visual complexity can explain different types of relationships with beauty. Similarly, individual differences have also been identified as being able to explain observed discrepancies regarding the inverted U-shaped curve (Güçlütürk et al., 2016). The comparability across studies is generally rather low due to different types of sensory and artistic stimuli that have been selected on various grounds, different presentation durations, and a high variability of operationalizations of collative variables and measures of hedonic tone. We still know very little about the dynamics of aesthetic experiences (Muth et al., 2015). Most current studies and reviews (e.g., Chmiel & Schubert, 2018; Nadal et al., 2010), when examining the relationship between a collative variable and a measure of hedonic tone, do not assess the role of subjective and/or physiological arousal (Berlyne, 1974, refers to such scales as *internal-state scales*). In this doctoral thesis, subjective arousal will be measured, which is especially relevant when studying affective stimuli.

More generally, we argue for a more comprehensive research approach (e.g., in terms of the number of variables and stimuli studied and the methodologies employed) when aiming to contribute to theory building in the field of empirical aesthetics. Such an approach was promoted by the *New Experimental Aesthetics*, for example, by arguing that subjective ratings should comprise three types of scales, namely *descriptive scales*, *evaluative scales* and *internal-state scales* (Berlyne, 1974).

When studying how objects induce pleasure, Fechner differentiated between a *direct* and an *associative* factor, the former referring to structural object features, and the latter to individual differences due to the perceiver's learning history (i.e., knowledge, memories and past experiences) (Nadal & Urena, 2022). Berlyne (1974) was mostly interested in subjective responses to collative stimulus properties and their role in aesthetic responses, but he and his followers also studied individual differences. Importantly, Berlyne (1971) emphasized that it is the interaction between an individual and the stimulus properties (i.e., not the stimulus properties alone) that elicits exploratory behaviour. Studying both bottom-up and top-down factors determining aesthetic responses has had a long tradition, and thus these factors are also mentioned in the context of other frameworks and models of aesthetic experiences (Jacobsen, 2006; Leder & Nadal, 2014; Locher et al., 2010; Pelowski, Markey, et al., 2017; Specker et al., 2021). Individual differences are also acknowledged in the studies of this doctoral thesis by assessing the familiarity with a stimulus and by studying the role of personality and interest in aesthetic responses.

Regarding bottom-up factors, the study of *complexity* has had a long history in psychology (Birkhoff, 1933; Eysenck, 1942), probably because of its evolutionary significance in humans' relationship with their natural environment (e.g., Godfrey-Smith, 1996; Kaplan & Kaplan, 1989) and its link with curiosity and motivational behaviour (Berlyne, 1960; Sun & Firestone, 2021). Visual complexity has also been shown to influence the performance in a wide range of cognitive tasks, such as affective memory and perceptual learning (see Güçlütürk et al., 2018). Similarly, the auditory complexity of messages has an impact on memory and arousal (Potter & Choi, 2006). In the context of information theory, stimulus complexity is related to uncertainty due to the higher information-processing load in comparison to processing simple stimuli (e.g., de Fleurian et al., 2017; Donderi, 2006; Van Geert & Wagemans, 2020). Berlyne and his followers were particularly interested in complexity because of its relation to preference, and early research already involved auditory (e.g., Crozier,

1974; Heyduk, 1975; McMullen, 1974) and visual (e.g., Day, 1967; Eisenman, 1966; Saklofske, 1975) stimuli, including the study of objective measures of complexity (e.g., Chipman, 1977; Vitz, 1966). Since then, the interest in studying complexity – within the context of Berlyne's theory and beyond – has not declined, and the sensory, cognitive and affective responses to unimodal stimuli varying in perceived complexity have been well documented (e.g., Bradley et al., 2011; Hagerhall et al., 2008; Nagle & Lavie, 2020; Sun & Firestone, 2021; for a review on complexity and aesthetic processing, see Marin, 2022). Moreover, the multidimensional nature of visual (Berlyne, 1960; Berlyne et al., 1968; Nadal et al., 2010) and musical (Mauch & Levy, 2011; Pressing, 1999; Streich, 2007) complexity has been acknowledged for a long time. However, prior to the beginning of this doctoral thesis, visual and musical complexity had rarely been discussed and investigated within a common framework (but see Boon et al., 2011).

Berlyne (1971, p. 43) theorized on how the communication between an artist and the perceiver through a work of art may work. He argued that a work of art consists of elements that can communicate information stemming from four types of sources: *semantic* ("external objects or characteristics of the world in general"), *expressive* ("information about processes within the creative artist"), *cultural* and *syntactic* (structural). Information from these sources can correlate and influence each other, but the sources can also communicate unique, independent information during the transmission process. Berlyne subsumed expressive, cultural and syntactic information under the term *aesthetic information* (see also Frank, 1959; Moles, 1958).

Most of the research in the tradition of Fechner and Berlyne has focused on the role of stimulus characteristics (i.e., formal properties and structural information) in aesthetic pleasure. However, at the end of the 19th century an alternative approach was already proposed by Vischer, who emphasized the idea of *Einfühlung* (feeling oneself into) during aesthetic experiences. In the 20th century, this idea was further developed by Lipps, who also influenced Bullough's idea of *psychical distance* (i.e., aesthetic distance) (for a review, see Cupchik, 2022). What all of these *projectionist* approaches have in common is a perspective on aesthetic experiences that emphasizes the interpretation and meaning of artworks, including emotional responses. Emotions have not been at the centre of empirical aesthetics for a long time (Tan, 2000), possibly due to the influence of Berlyne (see 1971, p. 62), who was skeptical about the concept of emotion and the dawn of emotion psychology (see also

Cupchik, 1988; Menninghaus et al., 2019). As a matter of fact, Konečni (1996) criticized Berlyne's disinterest in sympathetic arousal and emotion. But since the time of the *New Experimental Aesthetics*, emotion psychology has developed into a thriving independent research field. Silvia (2005) argued for an integration of knowledge gained in emotion psychology (i.e., appraisal theory) to the study of aesthetic responses, and further, that arousal alone cannot explain aesthetic responses (for a similar argument see Brown & Dissanayake, 2009). Indeed, emotion has received some attention in the field, for example, by studying emotions experienced in the museum (Tinio & Gartus, 2018) and by studying psychophysiological responses to visual art (e.g., Gerger et al., 2018; Pihko et al., 2011) and music (e.g., Bullack et al., 2018; Ogg et al., 2017). Cupchik (2016, 2022) has tried to reconcile the positions of formalists and projectionists by arguing that in the first moments of aesthetic processing the structural properties are in the foreground, and affective responses are a product of perceptual and cognitive processing (see also Craik & Lockhart, 1972). Then, when a stimulus is being perceived for a longer time, the perceiver's personal and cultural history is influencing the search for meaning. In a similar vein, Leder et al.'s (2004) processing stage model of aesthetic appreciation and aesthetic judgements incorporates both formalists' and projectionists' perspectives. Several recent neurobiological models of aesthetics and the arts focus on emotion, motivation and reward (Chatterjee & Vartanian, 2014; Rolls, 2017; Salimpoor et al., 2015), which could be interpreted as an alternative approach to formalist and projectionists models because they follow a motivational approach, which is rather in the tradition of Berlyne (1971).

In the current doctoral thesis, it was decided to work explicitly with stimuli of affective and – in the case of visual stimuli – semantic contents to gain a more comprehensive view of Berlyne's theory. Ecological variables have often been ignored in the study of collative motivation (e.g., Nadal et al., 2010), which is surprising because Berlyne himself started off as a motivational psychologist. As stated above, Berlyne was particularly interested in the study of collative variables in aesthetic experiences, but also acknowledged the role of psychophysical (e.g., intensity, brightness and colour) and ecological variables (i.e., signal value, associations with biological significant events such as threats, health and food). According to Berlyne (1971, 1974), all of these variables contribute to arousal changes in the perceiver, but collative variables are hypothesized to be the main determinants, therefore, they have been more fre-

quently studied by researchers. It is important to note that the various types of collative variables have been frequently studied in isolation (for a review, see Marin, 2022), and only a few studies have investigated collative variables together with psychophysical variables, such as complexity and colour in abstract paintings (Aslanoglu & Olguntürk, 2019). Markovic (2011) compared perceptual (color, form, space and complexity), semantic (e.g., expression, ideology) and affective (e.g., arousal, hedonic tone) judgements of representational and abstract artworks. His research was partly influenced by Berlyne's theory but mostly focused on differences between artistic styles. In the musical domain, Martindale and Moore's study (1989) deserves mentioning because they studied how semantic factors and collative properties determine preference for melodies, providing some counter-evidence for Berlyne's theory. In a similar vein, several studies used a wide range of visual stimuli (Martindale et al., 1990; Moore & West, 2012), for which collative and other types of variables were manipulated. These findings also support an alternative theory of preference based on meaningfulness and *prototypicality* (Martindale, 1984; Martindale & Moore, 1988; for criticism on prototypicality see North & Hargreaves, 2000). More recently, Chmiel and Schubert (2019) examined familiarity, complexity and a new variable called *unusualness* within the context of Berlyne's theory and music appreciation. Due to this limited number of studies incorporating several types of variables that communicate meaning through a work of art, this doctoral thesis follows a broader, more ecologically valid approach by considering psychophysical, collative as well as ecological variables when either selecting stimuli or running statistical analyses. Among other things, this approach also allowed us to compare the experience of environmental scenes with those of representational paintings, which offers insights into the nature of aesthetic experiences within a motivational framework.

Emotions are generally regarded as adaptively useful in promoting survival, and each emotion probably evolved to solve a specific adaptive problem (Tooby & Cosmides, 2008). Humans in the evolutionary past were confronted with conditions such as fighting, escaping predators, loss of status due to failure, or death of a family member. Specific emotion programs thus presumably evolved due to repeated encounters with these situations and enabled a dynamic interaction with the environment. Affective stimuli (in comparison to neutral stimuli) capture attention and benefit from enhanced perceptual processing. Induced emotions have been found to influence cognitive processes, and in a healthy population individual differences (sex,

personality and age) in affective processing have been described (for a review, see Dolcos et al., 2020). Affective responses to visual and musical stimuli can be simply described by *core affect* (Barrett & Bliss-Moreau, 2009; Russell, 2003; Russell & Barrett, 1999), which refers to a basic feeling, a blend of how good/bad and how energized/enervated one is feeling. Wundt already argued that momentary affective states can be described by pleasantness, arousal and intensity, which build a unified state (see Barrett & Bliss-Moreau, 2009). More formally, core affect is nowadays understood to consist of only two independent dimensions, namely pleasantness (valence) and arousal, as in Russell's (1980) influential *circumplex model of affect*. This 2-dimensional model can be used to describe all types of more complex emotions because core affect is thought to underlie and trigger an emotion episode after exposure to an affective stimulus. As a result of its simplicity and ease of use in experiments, this model has been frequently employed to study emotional responses to all sorts of visual (e.g., Fekete et al., 2023; Marin et al., 2012; Olofsson et al., 2008) and musical (e.g., Gingras et al., 2014; Nineuil et al., 2022; Ogg et al., 2017) stimuli. Some authors have argued for its superiority over other emotion models, such as the basic emotions paradigm in the domain of music (Cespedes-Guevara & Eerola, 2018; Eerola & Vuoskoski, 2011). Psychophysiological responses to musical and visual stimuli have also been directly compared, with similarities observed for the pleasantness dimension and differences for the arousal dimension (Kim & Wedell, 2016). Interestingly enough, it needs yet to be clarified whether the dimensions of arousal and pleasantness are dependent or independent when experiencing visual stimuli, both at the subjective (Kuppens et al., 2013; Yik et al., 2023) and neurophysiological level (Haj-Ali et al., 2020). It was decided to use Russell's circumplex model of affect as a framework to select affective visual and musical stimuli for this doctoral thesis. The main reasons for this choice were that the model is simple and has shown to be applicable to both the visual and musical domains. Moreover, arousal and pleasantness are core components of both the two-dimensional emotion model and Berlyne's theory.

The combined study of emotion (core affect) and complexity has been of interest in the field of general psychology for a while, for example, when studying the perception of environmental scenes and correlated eye movement patterns as well as neural responses (e.g., Bradley et al., 2007; Bradley et al., 2011; Kuniecki et al., 2017; Löw et al., 2013; Madan et al., 2018). Most of these studies suggest that affec-

tive processing may be largely independent from complexity processing in visual scene perception (i.e., independent main effects were observed). However, another line of research on scene perception suggests that affect has an impact on visual processing (Markovic et al., 2014; Mather et al., 2016; Pourtois et al., 2013), which makes studying affective scenes a timely topic. Affective contents (especially high arousal) has also been shown to facilitate cognitive processing (Schimmack, 2005; Schupp et al., 2003; Storbeck & Clore, 2008), and moreover, emotional and cognitive contributions to perceptual processing cannot be always clearly separated (Pessoa, 2008, 2022). Complexity, arousal and pleasantness are now common variables in stimulus sets of environmental scenes (Marchewka et al., 2014). Similarly, in the field of visual aesthetics, recent stimulus data sets are based on these, and other, variables of interest (Fekete et al., 2023). In this doctoral thesis, emotional arousal, valence (pleasantness) and complexity are central variables in all the presented research studies. One major goal is to investigate how these variables are related to each other and to compare these relationships across sensory/artistic domains.

The results will be of relevance for theoretical advances in the fields of empirical aesthetics and emotion research alike (Marin, 2015). Our findings can also be discussed within the context of aesthetic emotions (Juslin, 2013; Marković, 2012; Menninghaus et al., 2019). Researchers have proposed that there is a difference between emotions elicited by objects of motivational relevance and those elicited by aesthetic objects, such as music, visual art and so on (Scherer, 2004). Moreover, certain complex emotions, such as nostalgia and awe, have been discussed to be typical of art experiences (e.g., Barrett et al., 2010; Takano & Nomura, 2022). However, the existence of aesthetic emotions (in the sense of emotions of a specific quality that differs from everyday emotions) has been recently debated in the field (e.g., Menninghaus et al., 2020; Skov & Nadal, 2020), which demonstrated the timeliness of the topic.

A comparative approach to empirical aesthetics may also highlight the key differences between aesthetic experiences and everyday experiences (Marin, 2015). There has been a long-standing philosophical debate about the nature of aesthetic experiences (Cupchik, 2002; Hayn-Leichsenring & Chatterjee, 2019). How do human-made art and other objects part of the natural world induce pleasure and aesthetic experiences in humans? What makes aesthetic experiences so special? It is difficult to develop psychological research paradigms to address these questions

(Baltissen & Ostermann, 1998; Gerger et al., 2014), and the field of neuroaesthetics has also only partly provided satisfactory answers (Chatterjee & Vartanian, 2016; Nadal & Chatterjee, 2019). Besides the proposal to study the dynamics of aesthetic experiences in comparison to everyday scene perception (Marin & Leder, 2016), we propose to study eye movements and natural viewing behaviour when viewing environmental scenes vs. representational paintings (Marin & Leder, 2022). In addition to examining Berlyne's theory from a comparative perspective, these methodologies grounded in experimental psychology may contribute to our understanding of the nature of aesthetic experiences.

Since the inception of empirical aesthetics in the second half of the 19th century, researchers have studied aesthetic responses to a wide range of visual, musical and literary stimuli, to name a few (for a review, see Nadal & Ureña, 2021). Even in the early days of empirical aesthetics, some researchers already showed an interest in studying more than one type of aesthetic object or artform, such as Kate Hevner (1898-1984). Mostly known for her contributions to the study of emotions expressed by music (Hevner, 1936), she also conducted research involving poetry and paintings (Hevner, 1935, 1937). Later, Berlyne's movement of the *New Experimental Aesthetics* (1974) studied responses to visual and musical stimuli under one common theoretical framework, but rarely within one study led by the same authors (e.g., Hare, 1974).

When this doctoral thesis was started in 2012, studies directly comparing aesthetic responses to two types of stimuli of the same modality, different modalities or art forms were rare (Augustin et al., 2012; Baltissen & Ostermann, 1998; Boon et al., 2011; Vessel & Rubin, 2010). This limited the direct comparison of results across sensory and artistic domains due to issues related to stimulus selection, background variables of participants, analysis methods and so on. Therefore, back then we initiated a research program that has promoted the *systematic comparison of aesthetic experiences across sensory domains and art forms* to facilitate theory building in the field (Marin, 2015; Marin & Leder, 2013). This is in the tradition of Berlyne (1974) who was interested in the study of more than one artform. Here, the term *aesthetic experience* refers to a person's subjective experience and an understanding of the processing of its physical, situational, stimulus- and person-related determinants (de Cruys & Wagemans, 2011; Leder et al., 2004; Leder & Nadal, 2014). Understanding factors affecting aesthetic experiences goes alongside the study of their underlying

neurophysiological foundations (Chatterjee & Vartanian, 2014; Jacobsen & Beudt, 2017; Pearce et al., 2016; Pelowski, Markey, et al., 2017; Strijbosch et al., 2022; van Leeuwen et al., 2022). Therefore, in the long run, collaboration between experimental psychologists and neuroscientists is indispensable to move forward in the field (see also Brown, 2022).

Brown (2022, p. 23), a neuroscientist not particularly focusing on aesthetic experiences but rather on the nature of human art forms, similarly argues that "a comparative analysis of the arts provides greater insight into each artform than is possible by looking at art forms in isolation." He suggests classifying the arts into *narrative arts* (function to tell stories) and *coordinative arts* (function to synchronize action). More specifically, he characterizes the visual arts by static narrativity, while the coordination of voices is typical for music. This differentiation also alludes to the possibility to classify visual arts as *static* and music as *dynamic*, which also relates to the way how information is communicated by the two art forms. Brown (2022, p. 21) further argues that in the narrative arts people are viewed as "*characters* who reside in fictional storyworlds, whereas in the coordinative arts, we relate to people as *partners* who exist in our own real world." Brown (2022, p. 73) also states that the "arts are more than aesthetics", suggesting that we should not only focus on responses to aesthetic objects, and he particularly emphasizes the social communication of emotion, which in his view underlies artistic communication due to its expressive nature. He regards aesthetic emotions (i.e., "valenced reactions to the appeal of objects", such as pleasure, displeasure, attraction, repulsion, liking and dislike, Brown, 2022, p. 80) as only one type of emotion in a broader spectrum of emotions that people experience when interacting with or producing art. In his discussion about emotions and the arts he also reminds us of the important differentiation between emotions that are *expressed* by an artwork and those that are *felt* by the perceiver (see also Schubert, 2013; Tan, 2000). For instance, a sad musical piece can lead to different types of felt emotions in the listener, with sadness being only one out of several possible reactions (Herdson et al., 2023). This may also hold for visual arts. Brown's awareness of the role of emotions in aesthetic communication is mirrored in our approach to study affective stimuli across sensory and artistic domains.

The studies included in this doctoral thesis involve three types of stimulus categories (Table 1), varying in their aesthetic appeal and/or sensory modality, namely affective environmental scenes, affective representational paintings and in-

strumental music. The comparison between environmental scenes and representational paintings allows us to draw conclusions about differences and similarities between everyday visual experiences of motivational relevance and visual aesthetic experiences (e.g., Gerger et al., 2014). Both stimulus types contain semantic and affective meaning, which enables a valid comparison. As expressed above, we decided to study affective visual scenes because they determine most of our everyday visual experiences and thus constitute an excellent control stimulus category. Representational paintings have been the prevailing visual artform until the beginning of the 20th century although abstract and geometrical forms of rock art are the earliest manifestations of visual art (Henshilwood et al., 2011; Texier et al., 2010). We included music as a third stimulus category because music is, besides visual art, one of the most frequently studied aesthetic stimuli and of evolutionary significance (Brown, 2022). Therefore, we hoped to make a valuable contribution to the field by comparing visual and auditory aesthetic responses. Moreover, all types of stimuli vary naturally in terms of their visual/musical complexity. Among other things, our studies will offer insights into whether the concept of complexity can be similarly understood across sensory domains and how concepts such as arousal, pleasantness and complexity are related to each other in each stimulus category, which will enrich theory-building (see also Marin, 2015). All our stimuli communicate affective contents in a natural way, and we thus followed an *analytic approach* (Berlyne, 1974). Studying affective stimuli across sensory domains within Berlyne's framework is timely because stimuli with strong affective contents and those containing human figures and human faces have often been avoided to minimize the influence of ecological variables (Nadal et al., 2010).

Table 1

Characteristics of the Three Types of Stimuli Used in this Doctoral Thesis

Environmental scenes	Representational paintings	Music
Visual modality	Visual modality	Auditory modality
Non-artistic	Artistic	Artistic
Semantic meaning	Semantic meaning	----
Affective meaning	Affective meaning	Affective meaning

The systematic and comparative study of both visual and musical responses is timely and meaningful from a psychological perspective (see also Brown, 2022) be-

cause music and visual art share several characteristics and social functions. Vocal and instrumental music are an essential part of human social life in all cultures (Blacking, 1973), which is also the case for visual art (Davies, 2012). The origins of music(ality) are highly debated (Ravignani, 2018; Wallin et al., 2000), which also holds true for the origins of visual art (Dissanayake, 2019; McDermott, 2021; Straffon, 2019), but interestingly, similar theories, such as those related to sexual selection and social cohesion, have driven these debates. Moreover, most researchers agree that music is a means of social and emotional communication (Juslin & Sloboda, 2010; Schäfer et al., 2013), which affects human's behaviour, health and well-being (Zaatar et al., 2024). Similarly, emotional responses to visual art are very common (Silvia, 2005), and the psychological benefits of visual art experience have been well documented (Trupp et al., 2023). Unsurprisingly, then, both music and visual art have been used in a therapeutic context for a long time (Jensen & Bonde, 2018; Kaimal & Arslanbek, 2020; Li et al., 2021).

The comparison between music and visual art has had an old (philosophical) tradition, dating back to ancient Greek ideas. Among these ideas is the concept of the *Music of the Spheres*, which refers to how celestial bodies move in the universe in harmony and according to mathematical proportions. Plato relates the cosmos to colours and musical pitch in his *Republic*. Aristotle argued that colours can be produced by different proportions of black and white, and Pythagoras found that musical pitch is associated with the length of a vibrating string and that intervals are determined by ratios. The notion that proportions play a role for different qualities of colour and pitch as well as the concept of the *Music of the Spheres* lay the foundation for a comparison between colour and musical pitch, and then later, more generally for the one between music and paintings (Lill, 2012). Notably, music-colour synaesthesia is a phenomenon that is frequently studied among current psychologist and neuroscientists (Curwen, 2018). Moreover, abstract artists, such as Wassily Kandinsky (1866-1944) and Paul Klee (1879-1940), showed a strong interest in relating features of music to composition and expressionism in abstract art (McDonnell, 2007).

The critic and composer Charles Avison (1709-1770) compared music with paintings by discussing several parallels and differences. One analogy that was frequently referred to in later times was Avison's suggestion that light and shade in paintings can be compared to concord and discord in music (Schueller, 1953). Interestingly, Daniel Webb noted that music and paintings differ in their ability to commu-

nicate passions, with music being superior. He wrote that music catches “the movements of passion […] as they spring from the soul; paintings wait until they rise into action, or determine in character” (see Schueller, 1953, p. 339). This quote refers to music’s implied temporal nature and the feeling of movement when listening to music (Damm et al., 2020; Etani et al., 2024), a phenomenon that clearly differentiates the perception of static visual art from music perception. As Stechow (1953, p. 329) points out, “[i]t is a truism that musical structure is organized time, while pictorial structure is organized surface.” When listening to music, musical information is expanding over time (Levitin et al., 2018), whereas when looking at a novel painting, or any other static visual object, the gist of the pictorial scene is quickly represented in perceivers (Anderson et al., 2022), and if interest remains after the basic scene interpretation, a more detailed elaboration of the visual information will follow (Locher, 2015; Sanocki et al., 2023; Schwabe et al., 2018). In this sense, and as shown by eye movement studies (Goller et al., 2019; Henderson, 2003; Kowler, 2011), art perception is an active, dynamic process (Augustin et al., 2011; Verhavert et al., 2018) somewhat comparable to music processing; however, it clearly differs from music perception due to its temporal structure, and thus its dynamic nature of unfolding new musical information over time (Koelsch et al., 2019). This underlies the major difference in how humans apprehend music and static visual scenes, and it is the temporal structure that allows for phrase structure and syntax in music. As Brown states (2022, p. 121), “[w]hile one can certainly talk about ‘composition’ in visual art at the level of spatial organization and hierarchical structure (Arnheim, 1988; Mittelberg, 2013), there is no real basis for talking about phrase structure or syntax at the level of temporal organization.” Boon et al. (2011, p. 8) also refer to the distinct nature of music by referring to time: “The arrow of time is intrinsic to musical expression: music emerges from silence and returns to silence.”

Another difference between the processing of music and figurative images (i.e., environmental scenes and representational paintings) clearly refers to the semantic contents of images. In environmental scenes and representational paintings, semantic meaning is generated by the objects and figures depicted in a visual scene (Schepman & Rodway, 2021). Music, however, communicates semantic meaning differently (and to a much lesser degree) than figurative visual scenes do (Koelsch, 2011; Watt & Ash, 1998; Zhu & Meyers-Levy, 2005). Thus, one could argue that abstract art, lacking clear semantic contents, would be more comparable to music than

representational art. While there is some value to this argument, it seems that one needs to consider different artistic styles and music genres in greater depth. Actis-Grosso et al. (2017) studied how figurative and abstract art relate to classical music and jazz. They showed that classical music is more likely to be judged as figurative, whereas jazz is more likely to be judged as abstract, which suggests a structural correspondence between artistic styles and music genres. Therefore, our choice of studying classical music seems to be appropriate from a structural perspective. Moreover, we aimed at studying representational paintings because we considered the comparison between environmental scenes and representational paintings as more important in terms of theory building than the comparison between music and abstract art. Another reason for our choice of stimuli was that the systematic comparison between environmental scenes and representational paintings allowed us to gain insights into the nature of visual aesthetic experiences.

Empirical research on comparing aesthetic responses across the arts was scarce at the time when this doctoral thesis was initiated. Augustin et al. (2012) had investigated the repertoire of verbal descriptors of aesthetic experiences in response to music, visual art and film. The authors found similarities and differences in word usage across the three types of art forms. For example, emotion-related words were used more frequently in the context of music and film compared to visual art. Brown et al. (2011) presented a meta-analysis of neuroimaging studies, involving either pleasantness, attractiveness or liking ratings, across the four major sensory modalities. They reported activation of the right interior insula across all modalities, suggesting that the appraisal of valence of objects lies at the heart of aesthetic experiences. Boon et al. (2011) analyzed and discussed objective complexity (information theoretic measures), subjective complexity and social complexity of a small number of works by Jackson Pollock and Johann Sebastian Bach, offering fruitful ideas for a comparison of complexity across artistic domains.

Not only is the comparison between music and the visual arts meaningful in order to get a more comprehensive view of aesthetic experiences, but the comparison between the cognitive and affective processing of music and environmental scenes is also timely. In our everyday lives, sensory information from different modalities is usually not encountered in isolation. For example, visual, auditory, or haptic information inputs can follow each other (cf. *crossmodal interaction*) or occur simultaneously (cf. *multisensory integration*) in situations when a human needs to judge

whether an animal is going to attack, a car is approaching, or when a mother tries to soothe her crying baby. The ability to process these information inputs effortlessly and effectively is a remarkable evolutionary achievement and possibly represents one of the pillars of humans' and animals' survival in a multisensory environment (Nava et al., 2024). In fact, there is a wide variety of recent behavioural and neurophysiological evidence against the notion of strict modularity of sensory modalities in humans (e.g., Bravo et al., 2024; Gerdes et al., 2014; Shams & Kim, 2010) and animals (Munoz & Blumstein, 2012), which challenges the long-standing view of studying perception (and action) primarily from a unimodal and modular perspective (see also Barrett & Kurzban, 2006). Another ongoing line of research has reported that music listening can evoke images in listeners, namely visual experiences that relate to associations and even to storytelling (e.g., Hashim et al., 2023).

Before the advent of recording technology, music experience was characterized by a tight coupling of musical and visual signals because music co-occurred with visual information deriving from the musicians, the listeners, and in many cases, the dancers. More specifically, in the performing arts it is common to be exposed to a combination of affective cues stemming from different modalities, such as in music performance (e.g., Czepiel et al., 2023), dance (e.g., Christensen et al., 2014) or opera (e.g., Scherer et al., 2019). Moreover, in our modern world people often listen to music while pursuing activities such as driving, doing sports or walking through a city, and they are commonly exposed to background music in restaurants, shops or public places (North & Hargreaves, 2008). In short, systematically comparing sensory and cognitive unimodal processes across domains (i.e., environmental scenes vs. music) may be an initial and vital step in enhancing our understanding of how these sensory inputs may interact in crossmodal and multimodal settings. The implications of our results are manifold and will be of interest not only to those who study the cognitive specificities of visual and musical aesthetics, but also to those who study multimodal perception, such as in music performance (e.g., Jakubowski et al., 2020), dance (e.g., Christensen & Calvo-Merino, 2013), film (e.g., Tan et al., 2013), website design and virtual art/music experiences.

The presentation of the current research context within the field of empirical aesthetics has revealed that while there has been a great deal of research focusing on a single artform (e.g., paintings or music), remarkably few studies have been designed to address relevant questions from a broader perspective. This

(over)specialization is surprising because researchers like Berlyne (1974) already pursued a more comprehensive research program. However, it still remains to be seen whether a systematic comparative approach to empirical aesthetics can meaningfully contribute to theory building and the development of the field more generally in the long run (Marin, 2015). Due to the “reproducibility crisis” (Klein, 2014), the field of psychology has been facing serious problems regarding theory construction for decades (Borsboom et al., 2021). One possible way to tackle this issue is to critically put to test scientific theories, especially those that have been frequently cited in the field, before completely abandon them. If specific aspects of a theory can be replicated and other parts be improved, then such theories have still value, especially if they are of an explanatory character. It is then up to the researchers to either integrate the still valid parts of an otherwise outdated theory into other current theories, or to build a new, adapted theory centred on the core findings of the old theory.

1.2 Research questions and methods

This doctoral thesis addresses the issue of whether a comparative approach to empirical aesthetics is a fruitful path to make theoretical advancements in the field. More specifically, we aimed to study (mostly) subjective responses to three types of affective stimuli, namely environmental scenes, representational paintings and music, by using a set of selected variables (i.e., familiarity, complexity, arousal and pleasantness) relevant in the context of Berlyne’s (1971) psychobiological model. Therefore, the empirical research papers of this research program (Marin & Leder, 2013, 2016, 2018, 2022; Marin et al., 2016) comprise the same stimulus materials, with the aim of obtaining a more comprehensive view on the theoretical issues and discrepant research results surrounding Berlyne’s theory. We are aware of the role of individual differences in aesthetic responses (familiarity with stimulus materials, art/music interest, gender and personality traits) and thus either controlled for these influences in our statistical models or explicitly examined factors underlying individual differences.

By following a systematic comparative approach, our findings will also deepen our understanding of the nature of aesthetic experiences and aesthetic emotions, especially due to the comparison of responses to environmental scenes (non-aesthetic motivational visual stimuli) and representational paintings (visual aesthetic stimuli). If the aesthetic experience of representational paintings differs from that of environmental scenes, this should become obvious when simply studying the rela-

tionship between specific variables, the dynamics of aesthetic experiences, or when just studying how people are looking at these stimulus categories considering the role of the images' complexity and affective contents. Moreover, insights into the nature of visual and auditory aesthetic experiences will also be gained by directly comparing the perception of two common art forms grounded in different sensory modalities, namely representational paintings and music.

The selected core subjective measures comprise complexity (descriptive scale), arousal (internal-state scale), pleasantness (evaluative scale) and familiarity (descriptive scale to control for individual differences). We extended the range of evaluative scales to liking and beauty in Marin et al. (2016). Other frequently studied collative variables, such as novelty, and other dependent variables, such as interest(ingness) and meaningfulness, were not explored. Except for one study (Marin & Leder, 2022), we did not employ behavioural measures; instead, we used a wide range of computational measures of complexity and related them to subjective responses. Considering the restricted possible duration of our experiments, we needed to make a trade-off between the number of stimuli and the number of rating scales. To increase the generalizability of our results, we decided to work with large stimulus sets and a restricted number of rating scales. Most of the stimulus materials (musical excerpts and photographs of paintings) were especially selected and prepared for the purposes of this doctoral thesis.

We attempted to shed some light on the following specific research questions:

- How does the distribution of images/excerpts in the two-dimensional emotion space spanned by arousal and pleasantness differ across environmental scenes, representational paintings and music?
- How does the frequently reported positive linear relationship between complexity and arousal (controlled for familiarity) vary between the perception of affective environmental scenes, representational paintings and music?
- Can the inverted U-shaped curve between complexity and pleasantness be observed across the three affective stimulus sets?
- Does subjective arousal act as a mediator in the complexity-hedonic tone relationship in all three affective stimulus sets when familiarity is controlled for?

- Do concepts such as pleasantness, liking and beauty, subsumed under the term *hedonic tone*, relate similarly to complexity and arousal across the three affective stimulus sets?
- How does the performance of objective measures of visual complexity depend on the stimulus type (environmental scenes vs. representational paintings)?
- Can objective measures of visual complexity (based on Kolmogorov complexity and edge detection) in their analogous form be applied to the study of musical complexity?
- How does presentation duration (1 s, 5 s and 25 s) affect the relationship between objective and subjective complexity in the visual domain?
- When controlling for familiarity, how does presentation duration (1 s, 5 s and 25 s) affect the relationships between complexity, pleasantness and arousal in the visual domain?
- How does presentation duration (1 s, 5 s and 25 s) affect mean ratings of complexity, pleasantness and arousal in the visual domain?
- When controlling for familiarity, how do personality traits related to affect and art/music interest explain individual differences regarding the processing of complexity and arousal when perceiving environmental scenes, paintings and music?
- How does gender affect the relationships between familiarity, complexity, arousal and pleasantness, respectively?
- How does viewing behaviour differ between the perception of affective environmental scenes and representational paintings?
- How do arousal, pleasantness and complexity affect viewing behaviour? Is there an effect of presentation duration (long vs. short)?

The answers to these research questions will have profound theoretical implications and reveal whether parts of Berlyne's theory (1971) are still valid, and therefore, cannot be ignored in modern theory-building. Furthermore, the results will provide insights into whether aesthetic theories should rather focus on a single artform, or whether it may be worthwhile developing theories that have power to explain aesthetic experiences across art forms and sensory domains, which would be in the tradition of Berlyne and his *New Experimental Aesthetics*.

Regarding the research methodology, all the details are given in the respective research articles, so a brief overview should suffice to illustrate our approach. In general, we employed data collection and data analysis methods that are typical for the field of experimental psychology, and all studies were conducted in the laboratory under controlled conditions. As mentioned above, we focused on subjective ratings because we were mostly interested in the study of aesthetic experiences. These ratings were analyzed with standard parametric statistical methods, such as correlations and regressions, as well as more advanced mediation analysis. Following Berlyne's approach and due to an increasing interest in objective measures of complexity (e.g., Forsythe et al., 2011), we also studied objective visual and musical complexity. For this purpose, we used a wide range of conventional image statistics and compression algorithms for the analysis of the visual stimuli. Similarly, we used simple sound analysis algorithms and compression algorithms to determine musical complexity. To address questions regarding personality and individual differences, we employed several standardized scales and linear mixed-effects modelling to analyze the data. In the study exploring natural viewing behaviour, we measured eye movements with modern eye-tracking devices using infrared cameras in the laboratory under highly controlled conditions. In short, we employed a variety of research methods commonly used in the field of empirical aesthetics.

2. Publications

2.1 Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music (Marin & Leder, 2013)

2.1.1 Introduction

There is limited research investigating aesthetic processing across sensory domains within one common framework; therefore, this study (Marin & Leder, 2013, comprising four experiments), inspired by Berlyne's psychobiological model (1971), explored subjective responses to affective environmental scenes, paintings and music in a systematic way. Specifically, the comparative approach is reflected by studying two art forms of different modalities (representational paintings and music), as well as visual stimulus categories of low vs. high aesthetic quality (environmental scenes vs. representational paintings) with similar semantic contents. The main aim of this study was to gain a better understanding of how familiarity, complexity, arousal and pleasantness (key variables in Berlyne's model) are related in four affective stimulus sets varying in modality and aesthetic quality. By controlling for familiarity effects, this study also informed us about whether complexity and arousal are related in a similar way across stimulus sets, and whether arousal is a mediator in the complexity-pleasantness relationship. In recent years, arousal has often been ignored when investigating the relationship between collative variables and measures of hedonic tone, and this study filled this gap. Due to the fact that affective stimuli were employed in the study, possible gender effects were also considered in the analysis.

Another aim was to compare the performance of a large set of objective measures of complexity when analyzing environmental scenes and representational paintings. In the visual domain, objective measures of complexity based on compressed file size and edge detection algorithms have gained some attention in the field (Forsythe et al., 2011). Here, we explored a wide range of objective complexity measures and combined those that were previously used with measures that have been less frequently studied in empirical aesthetics at that time. In the musical domain, however, musical complexity has been somewhat neglected, especially the study of objective measures. Thus, the novel idea was to test whether objective measures of complexity that have been shown to be successful in the visual domain could be applied to the musical domain in their analogous forms. In this regard, studying compressed file sizes was considered a promising avenue. The idea to

combine subjective with objective measures is in the tradition of Berlyne's (1974) *New Experimental Aesthetics*.

We also sought to create stimulus sets that can be used for other empirical studies which are part of this doctoral thesis. For example, when pre-selecting stimuli, we considered the multidimensionality of complexity and focused on the number and variety of elements as the main factor underlying visual complexity (see also Marin & Leder, 2016). Divergent research findings regarding the relationship between complexity and hedonic tone may be due to complexity's multidimensional nature (Nadal et al., 2010). Visual stimuli in Experiment 1 and 2 were thus pre-selected to vary in complexity by including figure-ground compositions and complex scenes, arousal and pleasantness. For environmental scenes, which were IAPS pictures (Lang, Bradley & Cuthbert, 2005), standard ratings provided some guidance. In the musical domain, complexity varied naturally in Experiment 3 (solo piano music), whereas Experiment 4 mixed two types of musical styles of the same musical period (solo piano music and piano trios), which differed in the number of instruments and thus auditory streams. All stimuli were pre-selected by the author of this doctoral thesis, who is also trained as a musicologist.

We chose to study affective stimuli, especially since music is known to be a strong elicitor of emotions (Juslin & Sloboda, 2010). A meaningful comparison across stimuli of different modalities involving music is only possible if affective content is considered. Affective content (arousal and pleasantness) can also be easily varied if representational visual stimuli are employed. Moreover, our approach employed long stimulus presentation durations (25 s) across stimulus sets, ensuring that aesthetic experiences can be fully induced in participants (Smith & Smith, 2001).

2.1.2 Authorship contribution statement

Conceived and designed the experiments: MMM HL. Performed the experiments: MMM. Analyzed the data: MMM. Contributed reagents/materials/analysis tools: MMM HL. Wrote the manuscript: MMM HL.

2.1.3 Re-print of the published research article

OPEN ACCESS Freely available online

PLOS | ONE

Examining Complexity across Domains: Relating Subjective and Objective Measures of Affective Environmental Scenes, Paintings and Music

Manuela M. Marin*, **Helmut Leder**

Department of Basic Psychological Research and Research Methods, University of Vienna, Vienna, Austria

Abstract

Subjective complexity has been found to be related to hedonic measures of preference, pleasantness and beauty, but there is no consensus about the nature of this relationship in the visual and musical domains. Moreover, the affective content of stimuli has been largely neglected so far in the study of complexity but is crucial in many everyday contexts and in aesthetic experiences. We thus propose a cross-domain approach that acknowledges the multidimensional nature of complexity and that uses a wide range of objective complexity measures combined with subjective ratings. In four experiments, we employed pictures of affective environmental scenes, representational paintings, and Romantic solo and chamber music excerpts. Stimuli were pre-selected to vary in emotional content (pleasantness and arousal) and complexity (low versus high number of elements). For each set of stimuli, in a between-subjects design, ratings of familiarity, complexity, pleasantness and arousal were obtained for a presentation time of 25 s from 152 participants. In line with Berlyne's collative-motivation model, statistical analyses controlling for familiarity revealed a positive relationship between subjective complexity and arousal, and the highest correlations were observed for musical stimuli. Evidence for a mediating role of arousal in the complexity-pleasantness relationship was demonstrated in all experiments, but was only significant for females with regard to music. The direction and strength of the linear relationship between complexity and pleasantness depended on the stimulus type and gender. For environmental scenes, the root mean square contrast measures and measures of compressed file size correlated best with subjective complexity, whereas only edge detection based on phase congruency yielded equivalent results for representational paintings. Measures of compressed file size and event density also showed positive correlations with complexity and arousal in music, which is relevant for the discussion on which aspects of complexity are domain-specific and which are domain-general.

Citation: Marin MM, Leder H (2013) Examining Complexity across Domains: Relating Subjective and Objective Measures of Affective Environmental Scenes, Paintings and Music. PLoS ONE 8(8): e72412. doi:10.1371/journal.pone.0072412

Editor: Randen Lee Patterson, UC Davis School of Medicine, United States of America

Received December 11, 2012; **Accepted** July 13, 2013; **Published** August 16, 2013

Copyright: © 2013 Marin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: manuela.marin@univie.ac.at

Introduction

Complexity has been widely studied in psychology and related disciplines because of its relevance to humans' relations with their environments [1,2]. These disciplines have focused on the behavioral outcomes of sensory, cognitive and affective responses to stimuli varying in perceived complexity [3,6]. Understanding the impact of such collative stimulus dimensions as complexity, uncertainty and novelty, on hedonic value (e.g., preference, pleasingness and beauty) has been of paramount importance in empirical aesthetics, which primarily aims to study aesthetic preferences, mostly owing to Daniel Berlyne's contributions [7,10]. Research in light of Berlyne's *New Experimental Aesthetics* has been largely motivated by the question of what determines humans' preferences for certain kinds of stimuli over others. Perceived complexity has been reliably identified as a key determinant of hedonic value [11], and has therefore been included in current frameworks for the study of aesthetic experiences [12,15].

Research on complexity in the field of empirical aesthetics has been marred by contradictory findings, which may be due to several theoretical and experimental shortcomings. For example, the multidimensionality of complexity [16] is a major issue that needs to be controlled for in research designs. In this study, we focused on the number and variety of elements present in a visual and auditory scene. This dimension was found to be the strongest determinant of subjective complexity, more than organization or symmetry [16,17]. Moreover, researchers have mostly focused on subtle manipulations of stimulus complexity and neglected the emotional contents of

stimuli. This has led to a restricted, ecologically invalid way of studying the impact of complexity in aesthetic experiences, given the crucial role of emotions in such experiences. In addition, it is still not known whether the relationship between complexity and hedonic value is domain-specific or domain-general.

Here, we thus studied the effects of stimulus complexity on pleasantness (taken as a measure of hedonic value) by comparatively investigating large sets of visual and musical stimuli within the context of Russell's circumplex model of affect [18]. Moreover, we also examined the relation between a set of algorithms (including both previously used algorithms and new ones) to measure objective complexity in two kinds of affective visual stimuli varying in aesthetic quality and subjectively rated complexity. Our findings revealed that their performance depends in part on the stimulus type. In addition, our findings demonstrated that analogous types of algorithms developed for the auditory domain can be fruitful for the study of musical complexity and its relation to arousal and pleasantness. Finally, our analyses accounted for gender effects reported in response to affective visual [19] and musical stimuli [20]. In summary, we introduced a comprehensive approach to the study of subjective complexity, which not only comprised recent developments in measuring objective complexity but also crossmodal comparisons based on ecologically valid stimuli.

Berlyne's Collative-motivation Model and Divergent Research Findings

Berlyne's model predicts that people will generally prefer stimuli of intermediate complexity to simple and highly complex ones under normal arousal conditions [7,9]. This preference for intermediate levels of complexity gives rise to an optimal level of arousal, considered as an intervening variable, experienced as hedonically positive. The collative properties of a stimulus are, in Berlyne's model [9], the main determinants of arousal, though psychophysical (e.g., brightness, saturation, intensity) and ecological variables (e.g., innate or learned signal value, meaningfulness) also add to stimuli's arousal potential. Despite empirical evidence for an inverted U-relationship between a stimulus' perceived complexity and various measures of hedonic value in the visual [21,27] and musical domains [28,35], a considerable amount of counterevidence has also accumulated. These findings have either mostly revealed a linear relationship between perceived complexity and a specific measure of hedonic value or no clear relationship between these variables, both in the visual [16,36,41] and musical domains [42,47].

The discrepancy between findings of the relationship between perceived complexity and hedonic value in the visual and musical domains may not be explicable solely by the wide range of stimuli types employed in these experiments [16,44], which ranged from simple geometric forms and random shapes to artistic stimuli and landscapes in the visual domain, and from sequences of pure tones, melodies, chord progressions to music of different styles in the musical domain. Instead, the limited number of participants and stimuli used in several of the reported studies may have additionally affected the results. Moreover, the ecological validity of these materials may also play a crucial role in explaining the divergent findings [16,44]. The use of highly artificial stimuli, possibly due to the fact that a stringent manipulation of objective complexity has been deemed necessary by most researchers, is rather surprising considering that Berlyne is largely regarded as a motivational theorist [48], whose theoretical framework is primarily concerned about how humans explore their environment and display curiosity [7]. In addition, the different conceptions of complexity within psychology [49] and the related measurements and manipulations thereof may constitute another important factor explaining the discrepancy of the current research findings [16].

There is abundant empirical evidence for a multidimensional nature of complexity of visual [16,17] and musical [50] materials. Subjective visual complexity has been shown to be determined by stimulus features such as the number of elements, their organization and symmetry [16]. In a meta-analysis of studies exploring the relationship between complexity and hedonic value based on different definitions of visual complexity, Nadal et al. [16] suggested that manipulations of complexity on the basis of the number of elements present in a stimulus yielded a linearly increasing relationship between complexity and beauty, whereas varying the organization of elements showed an U-shaped or descending relationship, and manipulations along the dimension of symmetry an inverted U-shape relationship. However, regarding a comprehensive theory of aesthetic preference, other determinants besides complexity, such as stimulus prototypicality [51,52], expertise [44,53,55], familiarity [56,58], personality traits [59,62] and situational influences [63] may also have confounded the investigations on the relationship between subjective complexity, arousal and hedonic value.

Introducing Emotion to the Study of Complexity

The anomalous empirical findings in relation to Berlyne's collation-arousal model have led to serious criticism on theoretical grounds. For example, Berlyne's disregard of the dawn of cognitive psychology [64] made it impossible to integrate the interaction of thoughts and emotions in his framework of aesthetic experience. Another criticism concerns the disinterest in the diversity of emotions and their likely role in aesthetic experience. It is widely accepted that emotions play a fundamental role, for example, in the experience of visual art [65] and music [66], and thus it is not surprising that an affective approach to the study of aesthetic experience has been recently proposed [67,70]. For instance, the application of appraisal theories of emotion (for a review, see 71), which posit that subjective cognitive appraisals of events are the source of a wide range of possible emotions, may help account for the variability of emotions commonly experienced in response to artistic stimuli [72,73]. In contrast, Berlyne's collation-arousal model is restricted to arousal shifts and concomitant enjoyment, interest or aversion in response to a stimulus, yielding either simple positive or negative affective states. Another difference between Berlyne's theory and appraisal theories was pointed out by Silvia [70], who argues that

objective stimulus features and collative properties of a stimulus are not sufficient to create an emotion. Nevertheless, it needs to be stated that Berlyne's theory implies that the subjective experience of stimulus features modulates hedonic value, not the objectively measurable features of a stimulus [8,21]. In other words, with regard to subjective evaluations of stimulus features the two contrasting theories are not alien from each other.

Despite these initial efforts to introduce emotion to the study of aesthetic experiences induced by visual and auditory stimuli, previous research on complexity did not use stimuli that are strongly emotionally expressive, and even explicitly avoided using them [16,74]. To be specific, the stimuli under investigation were not selected within the context of current emotion models or on the basis of their degree of emotionality, but primarily on the basis of their varying degree of complexity [6,75,76]. We consider this as an important shortcoming because it is known, for instance, that the emotional content of stimuli modulates even the early neural processing of visual features [77,78].

The concept of arousal may be helpful in bridging the gap between emotion psychology and the study of stimulus complexity in the field of empirical aesthetics, particularly since arousal is an essential dimension in several emotion models [9,18,79,80]. However, studies in the tradition of Berlyne usually do not consider arousal as an independent variable, and moreover they rarely collect subjective arousal ratings or physiological arousal measures (but see 31,37,38,81,82). In order to incorporate arousal conceptually and empirically in the study of complexity, we propose to examine complexity within the context of Russell's circumplex model of affect [18]. The model considers arousal and pleasantness as two independent dimensions of affect, defining a wide range of different emotions. This two-dimensional model is highly suitable for crossmodal comparisons due to its simplicity and wide application in the visual and musical domains [83].

Gender effects in emotional processing are commonly reported and their underlying causes are manifold [84]. Specifically, gender effects with regard to emotional processing of visual [19,85,88] and musical [20] materials have been widely documented. Furthermore, subjective complexity judgments of photographs of fruits and vegetables have recently been found to be affected by gender [89]. Research on the perception of odors also revealed effects of complexity on preference that differed in females and males [90,91]. Therefore, we decided to consider the gender of the participant in the current research design and provide results for both genders separately.

Emotional responses are usually studied within the context of short presentation times. However, there is evidence that visual art is often experienced for much longer. For example, in a museum context Smith and Smith [92] reported an average viewing time of 27.2 s. Consequently, in order to enhance the ecological validity of our results and to study real aesthetic episodes, we chose presentation durations of 25 s in all four experiments. This also allowed for a fair comparison between the visual and musical domains since music is a dynamic stimulus that unfolds over time. Presentation durations of approximately 30 s are commonly used in research on musical emotions [83,93] and thus seemed to be appropriate for the purposes of the current study.

Computational Measures of Complexity

Which computational methods capture variations in complexity in a way that might be representative of how humans perceive complexity? The study of subjective complexity and its relationship to aesthetic experience has profited from work in the field of digital image processing and recent advances in finding automated measures of visual complexity [74,94,95]. These developments go beyond the initial mathematical approach of calculating objective complexity by considering the number of elements (lines and angles) and their heterogeneity in an additive way [96,98]. Instead, the application of various image compression techniques [99,103] and edge detection algorithms [74,100,103,104] has proved to be a reliable correlate of subjective visual complexity. Importantly, these measures were not developed by computer scientists to predict subjective complexity in the first place, but were only currently successfully applied to various types of visual stimuli by psychologists. The ratio between the original and the compressed file sizes of marine electronic charts and radar images [101,105], icons [103], line drawings [99,103] environmental scenes [74,104] and a wide range of artistic works [74] have been shown to be positively correlated with ratings of subjective complexity. It needs to be stated plainly that these types of stimuli were not selected within the context of a specific emotion model. In addition, only Forsythe et al. [74] have used compressed file size as a measure of complexity to test Berlyne's inverted U-shape hypothesis so far.

In the few studies that applied compression algorithms as objective measures of complexity to images, the file sizes of JPEG (Joint Photographic Expert Group) and GIF (Graphics Interchange Format) compression formats have yielded moderate correlations with subjective complexity judgments. In these studies, the stimulus presentation times [74,95,101,102,104] varied from several seconds to an unlimited exposure time in sorting tasks, suggesting that the correlation between compressed file size and subjective complexity may be independent of presentation time. Furthermore, Forsythe et al. [102] found that familiarity does not interact with objective complexity as measured by compressed file size and perimeter detection, which makes it a more reliable measure than subjective complexity ratings which are usually influenced by familiarity with the stimuli [106,107].

Donderi [4] explains that the success of image compression techniques to predict subjective complexity can be understood within the context of Algorithmic Information Theory (AIT) (for a review, see 108,109) which combines information-theory with the theory of computation (for a review on the Turing machine see 110): "Algorithmic complexity is defined in terms of the length of the shortest algorithm in any programming language, which computes a particular binary string" [49], p. 6. In principle, data compression algorithms analyze the visual information of an image, as described by a bit string, in order to

compress it to the extent that makes a valid reproduction of the original input possible [4]. The size of the resulting compressed data file correlates positively with the complexity of the input image. Simple images contain more redundant information that can be represented by a shorter string of bits, yielding a smaller file size than more complex images.

Besides the use of data compression, various edge detection measures developed in the field of image statistics have been shown to be another reliable way of measuring holistic complexity in the visual domain [74,95,100,102,104]. Edge detection algorithms, such as perimeter detection and Canny edge detection, detect changes in intensity at an image's edges. The more edges an image has, the higher the level of perceived complexity is [102,103]. Perimeter detection, a contour-based and a global measure of shape [111], has specifically yielded reliable - though moderate - correlations with subjective complexity across different stimulus sets [74,102,103]. Alternative approaches to edge detection include the analysis of root mean square (RMS) contrast, the standard deviation of the pixel intensities, which was recently applied to a set of day- and nightscapes by Cavalcante et al. [104]. The presence of high-contrast features is indicated by a high mean RMS contrast value, yielding a positive correlation with subjective complexity. The correlations between different measures of RMS contrast and subjective complexity were stronger (r ~.60) than the one for JPEG file size (r = .36). Thus, applications of RMS contrast measures to other types of visual stimuli seem promising.

Music is another affective domain that is widely studied within experimental psychology. The study of subjective complexity and objective complexity measures, however, has hardly received any attention in the field of music psychology [112], although research on musical expectation and its relation to complexity is flourishing [76,113,114]. In contrast to several recent reports of a positive association between compressed file size and subjective complexity in the visual domain, we are unaware of any studies using compression algorithms and related file sizes to predict subjective complexity in the musical domain. For example, Streich [50] modeled subjective musical complexity on the basis of twelve predictors extracted from audio, but this model did not include compression file size. Streich's model comprised four measures relating to the dynamic and spatial properties of an audio excerpt, one measure of timbral complexity, three measures of tonal strength, as well as four measures of rhythmic complexity. More recently, Mauch and Levy [115] proposed to objectively measure musical complexity by means of a structural change algorithm applied to changes in harmony, rhythm and timbre. Results of an internet-based experiment indicated that around 61.4% of the listeners agreed with the automated analysis, which is similar to the performance of Streich's model [50]. Nevertheless, it was shown earlier that compression-based approaches to classify MIDI files were successful in differentiating between musical works of different periods and between solo piano music by different composers of different periods [116,117]. Based on these findings, and the results reported above on studies using data compression in the visual domain, it can be surmised that the ratio between the original and the compressed file size may also be a significant predictor of subjective complexity of musical materials.

Overview of the Present Experiments

We investigated subjective and objective complexity in the visual and musical domains in a series of four experiments. The main goals of our approach were threefold: to compare the relationships between subjective complexity, felt arousal and pleasantness (by controlling for effects of familiarity) in large sets of visual and musical stimuli selected on the basis of Russell's circumplex model of affect [18]; to relate subjective complexity to measures of objective complexity; and to gain further insights into the similarities and difference between the perception of environmental scenes and art images. Our approach is characterized by the following decisions: First, pleasantness was chosen as a measure of hedonic value and arousal was not only considered as an unmeasured intervening variable within the framework of Berlyne's collative-motivation model. Instead, variations in arousal were inherent in the design of our experiments because we aimed at selecting representative stimuli that covered the arousal-pleasantness emotion space of a particular stimulus type as much as possible. Second, complexity was manipulated by pre-selecting stimuli that varied in the number of elements present in a visual or musical scene. Third, in order to make valid comparisons between the perception of environmental scenes and visual art possible, both types of stimuli were chosen to contain similar semantic contents. Fourth, we also explored the mediating role of arousal in the complexity-pleasantness relationship in the context of Berlyne's theory [9]. For example, Vettehen et al. [118] recently employed mediation analysis to address this question in a study on the effect of sensationalism on liking of television news stories. In this relationship, arousal was identified as a mediator.

In the field of image statistics, direct comparisons between image properties of environmental scenes and visual art have become popular because they may offer insights into the nature of artistic stimuli [119,121]. We were following this trend by comparing the performance of the very same set of objective complexity measures on two stimulus sets varying in motivational relevance and artistic quality (Experiments 1 and 2). Specifically, we compared the performance of objective measures related to object recognition processes (i.e., edge detection measures) and measures that capture visual information in a more abstract way (i.e., compressed file size) as approximations to human subjective ratings. In this study, we were particularly interested in the practical means of measuring objective complexity for the purpose of stimulus selection. In line with this, we neither primarily aimed at a comprehensive model of subjective complexity based on objective measures nor at an in-depth discussion of the performance of each objective measure with regard to aspects of perceptual and cognitive processing of complexity.

The goals of Experiments 1 and 2 were to compare the inter-relationships between subjective ratings of familiarity, complexity, pleasantness and arousal of affective environmental scenes and representational paintings, and further, to compare the performance of the very same set of

objective measures of complexity. In Experiment 1, stimuli consisted of affective environmental scenes selected from the International Affective Picture System (IAPS) developed by Lang, Bradley, and Cuthbert [122], a stimuli database widely used for visual emotion induction. In Experiment 2, we studied pre-selected affective representational paintings with similar semantic contents as the pictures used in Experiment 1. Since most previous studies only focused on a few measures of objective complexity, we aimed at assessing the performance of a wide range of measures previously reported to be successful in predicting subjective complexity (i.e., JPEG and GIF compression file size, perimeter detection, Canny edge detection, RMS contrast). Moreover, we extended this set of measures by including several potentially new measures of objective complexity, such as PNG (Portable Network Graphics) and TIFF (Tagged Image File Format) compression formats, measures of edge detection based on phase congruency [123], and the entropy of the image intensity histogram of a grayscale image [124]. It is also important to note that since it was already demonstrated earlier that compression file size is an indicator of subjective complexity [74,105], we selected pictures in JPEG format rather than uncompressed pictures as a starting point for all further transformations and analyses. This approach extends the implications and applicability of the current research findings since JPEG pictures are more easily accessible to the research community compared to (scans of) pictures in uncompressed formats.

For Experiments 1 and 2, we hypothesized that subjective complexity and arousal would be positively associated in both types of visual stimuli [9,21]. Furthermore, based on research by Nadal et al. [16], we surmised that subjective complexity and pleasantness would be linearly and positively associated because we mainly varied complexity by the number of elements present in a visual scene. In other words, we did not predict an inverted U-shape relationship between these variables as proposed by Berlyne [9]. Nevertheless, we predicted that the effect of complexity on pleasantness would be mediated by arousal, as proposed by Berlyne [9]. We also hypothesized that JPEG and GIF compression file sizes [101,102,104], perimeter detection [74,103], Canny edge detection [102] and RMS contrast measures [104] would yield a moderate positive relationship with subjective complexity if the image size was held constant, and that these objective measures would not correlate with reported familiarity [102]. Last, we predicted that measures of objective complexity would correlate positively with pleasantness [74].

Another major goal of the study was to show the fruitfulness of a comparative approach to the study of emotion and complexity by demonstrating the application of objective measures of complexity to the musical domain (Experiments 3 and 4) by using analogous measures to those used in the visual domain (Experiments 1 and 2). As such, this paper contributes to a wider theoretical discussion on the relationships between subjective complexity, arousal and pleasantness (i.e., one measure of hedonic value), and the use of objective measures of complexity in the affective musical and visual domains. In Experiment 3, we investigated subjective responses to a stimulus set recently developed by Marin et al. [83]. These musical stimuli represented one musical genre (Romantic solo piano music of the 19th century) and were evaluated within the context of Russell's circumplex model of affect [18]. In Experiment 4, we additionally varied complexity in a similar vein as we did in the visual domain (in Experiments 1 and 2) by changing the number of instruments audible in the musical excerpts, i.e., one versus three instruments. For this purpose, half of the stimuli were selected based on results of Experiment 3, and the other half consisted of piano trio excerpts of the same musical period. The hypotheses with regard to the inter-relationships between subjective complexity, pleasantness and arousal were identical to those of Experiments 1 and 2.

Since no study has reported on the use of compression file size as a measure of subjective complexity in the musical domain, we applied various audio compression formats to uncompressed WAV (Waveform Audio File Format) files in a first step. The auditory domain is prone to such an investigation because compression of audio files is a common and easily accessible tool. Moreover, analyses of acoustic and musical properties of audio signals relevant to subjective complexity have been made possible by recent developments in the field of music information retrieval [50,115]. Therefore, in analogy to edge detection algorithms and their application to predict subjective complexity in the visual domain, it was decided to analyze the event density per second of each musical excerpt by means of the MIRtoolbox [125], a collection of functions written in Matlab. Event density analyzes the overall amount of simultaneous (melodic, harmonic and rhythmic) events in a musical excerpt that can be perceived by a musical listener. This measure was not included in the models of musical complexity developed by Streich [50] and Mauch and Levy [115]. However, it can be conjectured that musical excerpts containing fewer musical events are judged as less complex than those with a higher number of events, and similarly, that compressed file size is a predictor of subjective complexity as previously observed in the visual domain.

Experiment 1

Methods

Ethics Statement. All four experiments reported in this article were conducted in accordance with the Declaration of Helsinki (revised 1983) and local guidelines of the Faculty of Psychology, University of Vienna. According to the Austrian Universities Act 2002 (UG2002) which held at the time the study was carried out, only medical universities were required to appoint ethics committees for clinical tests, application of medical methods, and applied medical research. Therefore, no ethical approval was required for the present study. Written informed consent was given by all participants who could withdraw at any time during the experiment without further consequences.

Participants. Thirty-six German-speaking psychology students (18 males, 23.9 ± 4.9 years, age range 21-42 years; 18 females, 22.6 ± 4.7 years, age range 20-41 years) participated in exchange for course credit. All participants had

normal or corrected-to-normal visual acuity. These participants did not take part in any other experiment of this study.

Materials. Pictures were chosen from the International Affective Picture System (IAPS) [122], a standardized picture system of colored photographs widely used in experimental psychology. Ninety-six colored pictures of realistic natural scenes were pre-selected on the basis of their standardized values of arousal and pleasantness to vary in hedonic content (high vs. low-arousing and unpleasant vs. pleasant). In addition, pictures were also chosen on the basis of their degree of visual complexity (figure-ground composition vs. complex scene), for which no standardized values were available. However, the classification of pictures into figure-ground compositions and complex scenes was guided by the results reported in studies on visual complexity and emotion using IAPS pictures [3,126]. Following Russell's circumplex model of affect [18], this pre-selection led to 24 pictures of specific hedonic contents (either low-arousing pleasant, high-arousing pleasant, low-arousing unpleasant, or high-arousing unpleasant). Half of the pictures were chosen to depict relatively simple figure-ground compositions (i.e., one figure with a uniform background), and the other half depicted complex scenes with several objects and a more varied background. The semantic content of these pictures varied largely, ranging from animals, plants, landscapes and food to human beings in everyday life scenes. Pictures showing erotic scenes, brand names and strong mutilation were not included in the stimulus set (see Supporting Information Stimulus List S1). All pictures were in landscape format (1024 x 768 pixels) and saved in their original JPEG format.

Procedure. Participants were tested either individually or in pairs separated by a wall in a quiet room. After signing the informed consent form, participants completed two practice trials in which they were familiarized with the task. Each trial was announced by a sentence appearing in the middle of the screen (Samsung SyncMaster S2443BW, 24-inch) and the picture followed after 5 s, displayed on a black background at a size of 33.90 x 25.50 cm. Participants, sitting 60-70 cm away from the screen, were instructed to look at the picture for the entire presentation duration of 25 s. After the picture had disappeared, the first out of four rating scales was displayed on the screen. This procedure was chosen to make possible a comparison between the visual and musical modalities because, due to the dynamic nature of music, global ratings could only be obtained after participants had listened to the excerpts. Participants indicated their choice by a mouse click and the next rating scale appeared immediately after. All pictures were rated on a seven-point Likert scale in terms of familiarity (from "unfamiliar" to "very familiar"), complexity (from "very simple" to "very complex"), pleasantness (from "very unpleasant" to "very pleasant") and arousal (from "very calm" to "very excited"). The order of the four ratings was the same for each participant and stimulus and the specific instructions said: "Please rate your familiarity with the contents of the picture," "Please rate your felt complexity of the picture," "Please rate the degree of pleasantness of your emotional experience," and "Please rate your felt arousal". From another recent experiment (unpublished data) involving the same stimuli sets as those used in the current Experiments 1 and 2, we found that the order of ratings (familiarity-complexity-pleasantness-arousal versus pleasantness-arousal-complexity-familiarity) did not affect the inter-relationships of the four variables under investigation. Importantly, participants were instructed to report their subjective feelings rather than the perceived emotional contents of the pictures (felt vs. perceived emotion, see 127) and to give all ratings spontaneously (although they were informed that there were no time constraints).

The pictures were presented in four blocks and the order of the blocks was randomized across participants. Within each block, 24 pictures of similar emotional content (either low-arousing pleasant, high-arousing pleasant, low-arousing unpleasant, or high-arousing unpleasant) but with varying complexity were randomly presented. This procedure was chosen in order to make it easier for participants to perceive subtle differences in terms of emotional contents and complexity. Participants were allowed to take a short self-timed break between these blocks. They were told that the four blocks would vary in terms of emotional contents. The entire experiment lasted around 75 minutes, after which the participants were debriefed and dismissed.

In order to assess objective measures of complexity using data compression, the original JPEG files were transformed into ZIP files (settings: maximum compression rate, compression mechanism deflate, dictionary size 32 kB, word size 64 and 2 CBU-threads) and 7z files (settings: maximum compression rate, compression mechanism LZMA, dictionary size 32 MB, word size 64 and 2 CBU-threads) using the 7-Zip file manager. The original JPEG files were also transformed into lossless GIF files using Adobe Photoshop CS5 (settings: palette local selective, colors 256, forced black-white colors, no transparency, dither diffusion 75%, exact colors and normal order of lines). Two other common lossless compression formats were included, namely PNG and TIFF. The TIFF files were compressed by using the Lempel-Ziv-Welch algorithm. All new file types were further compressed to ZIP and 7z files as described above, leading to twelve different compression file types in total. The file sizes of the pictures of the respective compression formats were assessed in bytes as given by Command Prompt.

Two conventional edge detection parameters were determined, namely perimeter detection and the Canny edge detection. Image analyses were conducted using the Image Processing Toolbox in Matlab (The MathWorks, Inc., Natick, Massachusetts, USA). In order to perform perimeter detection, grayscale images were transformed into binary images by considering the global image threshold using Otsu's method [128]. The perimeter function returns a binary image showing the perimeter pixels of objects, indicating the changes of intensity at the edges of an image. To be considered as a perimeter pixel, a pixel must have the value one (i.e., being white) in the binary image and be connected to at least one zero-valued pixel (i.e., being black). The default connectivity of four was used for the analysis. Four measures of perimeter detection were determined: the file size of the JPEG, PNG and TIFF compressions of the perimeter images conducted using the standard settings in Matlab and a raw measure based on

the sum of the white pixels representing the edges of an image after applying the perimeter function. The Canny edge detection method [129] is a widely used tool to detect weak edges appearing in combination with strong edges in grayscale images. The sensitivity threshold was automatically determined as well as the size of the Gaussian filter. JPEG, PNG and TIFF compression files were created from the images after the application of the Canny-algorithm, and a raw measure was determined by calculating the sum of the white pixels of an image.

Cavalcante et al. [104] showed that the RMS contrast of luminance values provide an alternative method for calculating edges in an image. A contrast map of a grayscale image reveals edges based on luminance contrasts, which proved to be a useful predictor of subjectively experienced complexity of a set of day- and nightscapes. We thus calculated the RMS contrast map of each IAPS picture in Matlab as an alternative way of detecting edges. RMS contrast does not depend on the spatial frequency content or the spatial distribution of contrast in the image and is defined as the standard deviation of the pixel intensities [130]. Around every pixel of an input image *I* a neighborhood of 15 x 15 pixels was considered in the calculation of the RMS contrast map **C** as

$$C\left(i,j\right)=\sqrt{1/MN\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left(Iij-\bar{I}\right)^2}$$

where intensities I_{ij} are the *i*-th *j*-th element of the two dimensional image of size *M* by *N*. $\bar{I}$ is the average intensity of all pixel values in the image. The image *I* is assumed to have pixel intensities normalized in the range [0, 1]. Following Cavalcante et al. [104], we calculated three measures of objective complexity based on the RMS contrast map: the mean of the RMS contrast values, the standard deviation of the mean RMS contrast values, and a measure α, which is a product of the mean and the standard deviation of the RMS contrast values.

Kovesi [123] developed another type of feature detection algorithm based on phase congruency. Whereas measures such as Canny edge detection are sensitive to variations in image illumination and blurring, phase congruency can be considered as an illumination and contrast invariant measure of feature significance. We calculated the maximum moment of phase congruency covariance (M), an indicator of edge strength, for each picture using the phasecong3.m function provided by Kovesi in Matlab (http://www.csse.uwa.edu.au/~pk/research/matlabfns/#phasecong). Similar to the RMS contrast measures, we computed three measures of objective complexity based on edge detection by phase congruency: the mean of the M values, the standard deviation of the mean M values, and a measure β, which constitutes the product of the mean and the standard deviation of the M values.

We also calculated the entropy of the image intensity histogram of a grayscale image [124], another potentially useful measure of objective complexity, which is included in the image processing toolbox of Matlab. Entropy refers to a statistical measure of randomness of an image: If all of the pixels have the same intensity value, the entropy of the image is zero. In other words, the higher the entropy value of an image, the larger its variation in intensity values.

Statistical Analysis. Statistical analyses were conducted in IBM SPSS Statistics version 19 (SPSS Inc., Chicago, IL, USA) or Matlab R2010b (The MathWorks, Inc., Natick, Massachusetts, USA). In order to address the issue of controlling for type 1 error in multiple testing of several correlations from the same matrix, we reported adjusted *p*-values calculated by following the sequential Bonferroni-Holm procedure [131]. For regression analyses, it was ensured that all assumptions (no multicollinearity between the predictors; independence, homoscedasticity and normality of the errors) were met for the variables in question. Mediation regression analyses were computed using the SPSS macro "MEDCURVE" [132]. All statistical tests were two-tailed at an alpha level of .05 if not otherwise indicated. This information refers to all four experiments of this study.

Results and Discussion

Subjective ratings for each picture were averaged across participants prior to any further analysis. In order to ensure that participants rated each stimulus in a consistent fashion, inter-rater reliability was assessed by computing the average measure intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency [133,134]. We observed a very high inter-rater reliability for all scales when both males and females were considered in the analysis, as evidenced by the following results: familiarity (ICC(2, k) = .93, 95% confidence interval (CI) [.94, .97]), complexity (ICC(2, k) = .94, 95% CI [.92, .96]), pleasantness (ICC(2, k) = .98, 95% CI [.98, .99]) and arousal (ICC(2, k) = .93, 95% CI [.91, .95]). The same type of analysis was also conducted for males and females separately, which revealed the following results for males: familiarity (ICC(2, k) = .92, 95% CI [.89, .94]), complexity (ICC(2, k) = .91, 95% CI [.87, .93]), pleasantness (ICC(2, k) = .96, 95% CI [.95, .97]) and arousal (ICC(2, k) = .88, 95% CI [.84, .91]). Similar results were obtained for the group of females: familiarity (ICC(2, k) = .92, 95% CI [.89, .94]), complexity (ICC(2, k) = .87, 95% CI [.83, .91]), pleasantness (ICC(2, k) = .97, 95% CI [.96, .98]) and arousal (ICC(2, k) = .87, 95% CI [.83, .91]).

An exploratory data analysis revealed that each of the variables referring to the different compression formats and other measures of objective complexity (edge detection, RMS contrast, entropy) contained several outliers that were 2 *SD* above or below the mean. A series of Shapiro-Wilk normality tests indicated that five variables of the measures of compressed file size, edge detection, RMS contrast and entropy deviated significantly from a normal distribution (all *p*s < .05), after removing the outliers. Furthermore, subjective ratings of familiarity, complexity, pleasantness and arousal were generally not normally distributed when all participants were considered together (all *p*s < .08), and when males (all *p*s < .020) and females (all *p*s < .23) were considered separately. Thus, it was decided to employ non-parametric analyses to investigate correlations between this set of variables in a first step. All relationships between the variables were visually inspected in order to ensure that the distributions of the stimuli

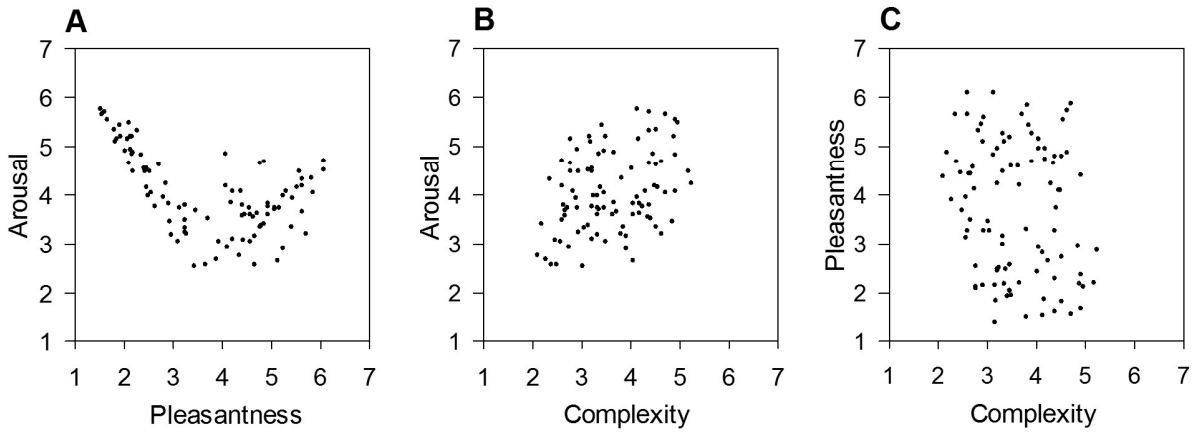

Figure 1. Relationships between pleasantness, arousal and complexity in a set of IAPS pictures. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Distribution of IAPS pictures in the pleasantness-arousal space based on mean ratings per picture. B) Relationship between mean complexity and arousal ratings. C) Relationship between mean complexity and pleasantness ratings.

doi: 10.1371/journal.pone.0072412.g001

followed a linear or monotonic function. Since further regression analyses were planned, values 2 *SD* above and below the mean were removed in all variables prior to these correlational analyses.

Note that we decided to present three types of relationships graphically: (a) the arousal-pleasantness relationship was considered as important because it shows the distribution of stimuli in the two-dimensional emotion space, which offers insights into the representativeness of the stimulus set, and (b) the arousal-complexity and (c) pleasantness-complexity relationships were considered as most relevant for the discussion of Berlyne's theory [9] in the context of emotional stimuli. In Figure 1a, the distribution of the 96 IAPS pictures in the arousal-pleasantness space is depicted. Arousal and pleasantness ratings followed a quadratic relationship, R^2 = .69, $F(2,91)$ = 99.19, p < .001, $y = -2.96x + .35x^2 + 9.54$. The pictures varied to a larger degree in terms of felt pleasantness (min. M = 1.39, max. M = 6.08, range = 4.69), nearly covering the complete seven-point scale, than in arousal (min. M = 2.53, max. M = 5.75, range = 3.22). Pleasant high-arousing pictures did not receive average ratings above five on the seven-point rating scale, which stood in contrast to a subset of 14 unpleasant high-arousing pictures. This finding of a smaller number of less arousing pleasant pictures can be explained by the fact that IAPS pictures with erotic scenes, generally inducing high pleasantness and arousal [122], were excluded in the pre-selection process of the current experiment. Another explanation refers to the often reported negativity bias in emotional processing [135], a bias in humans and animals to give greater weight to negative entities such as events and objects. In general, the current distribution of IAPS pictures for a long presentation time of 25 s resembles the distribution for a presentation time of 6 s found for the complete IAPS picture set [122]. Thus, the current stimuli can be considered as representative for this type of stimulus set.

Spearman's rank-order correlations were calculated between the four subjective ratings (familiarity, complexity, pleasantness and arousal) for all participants as well as for males and females separately (Table 1 and see also Figure S1) because gender differences in response to IAPS pictures have been reported earlier [19]. Familiarity correlated significantly with all other subjective measures when all participants were considered in the analyses. A significant negative correlation between familiarity and complexity was observed (r_s = -.31), indicating that more familiar pictures were rated as less complex. This finding is in line with earlier reports of effects of familiarity on subjective complexity ratings (for a review, see 102). Similarly, familiarity was moderately negatively associated with arousal (r_s = -.44): Unfamiliar pictures were experienced as more arousing than familiar ones. A further analysis of the difference between two Spearman rank correlation coefficients based on the Fisher r-to-z transformation [136] revealed that the correlation between familiarity and arousal was marginally stronger in males compared to females, z = 1.91, p = .056, but no other gender effects were observed. The positive relationship between familiarity and pleasantness was the strongest (r_s = .64), showing that familiar semantic contents of pictures was associated with a higher degree of pleasantness.

Two further associations in relation to complexity need to be noted. First, complexity correlated moderately positively with arousal (r_s = .36), in other words, more complex pictures induced higher degrees of arousal (Figure 1b). Next, we did not observe a significant association between complexity and pleasantness in the current data set (r_s = -.18), although a negative association between these two variables was visually present (Figure 1c). Environmental scenes of low complexity received higher pleasantness ratings. Furthermore, non-parametric partial Spearman's rank-order correlations were conducted to control for effects of familiarity in the associations

Table 1. Spearman's rank-order correlations between ratings of familiarity, complexity, pleasantness and arousal of IAPS pictures ($N_{All} \geq 94$, $n_{males} \geq 92$, $n_{females} \geq 92$).

Measure		Familiarity	Complexity
Complexity	f	-.30*	
	m	-.34*	
	All	-.31*	
Pleasantness	F	.66*	-.19
	M	.61*	-.17
	All	.64*	-.18
Arousal	F	-.27*	.30*
	M	-.51*	.37*
	All	-.44*	.36*

Note. *p < .05 after Bonferroni-Holm correction; f = females; m = males, All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

between complexity, pleasantness and arousal (Table S1). The relationship between complexity and pleasantness was considerably weakened (r_s = .01), whereas the positive relationship between complexity and arousal (r_s = .27) was still significant. The observed positive association between complexity and arousal is in line with Berlyne's theory [9,21]. However, we observed neither an inverted U-relationship between complexity and pleasantness, nor did we find support for a positive association between complexity and pleasantness as suggested by Nadal et al. [16] for non-artistic representational pictures. On the contrary, we found weak indications for a negative association between these variables when familiarity was not controlled for. This divergence may be due to several reasons: First, hedonic value was defined as beauty in Nadal et al. [16], whereas pleasantness was selected as a measure of hedonic value in the current study. Second, Nadal et al. [16] explicitly avoided the use of affective stimuli, while the current study selected stimuli of different affective contents. Third, Nadal et al. [16] did not explicitly control for familiarity effects when investigating the relationship between complexity and beauty, for example by obtaining subjective ratings for each picture.

In order to gain a deeper understanding of the underlying mechanism of the complexity-pleasantness relationship, a bootstrapping approach [132] was used to examine the existence of an indirect effect of complexity (independent variable X) on pleasantness (dependent variable Y) through arousal (mediator variable M), which would lend support to Berlyne's collative-motivation model [9]. Due to the nonlinear relation between arousal and pleasantness we conducted multiple regression analyses to compute the instantaneous indirect effect (Θ_x) by using the macro "MEDCURVE" for SPSS [132]. According to Hayes and Preacher [132] the instantaneous indirect effect "quantifies how much Y is changing at the point X = x indirectly through X's affect [sic] on M which, in turn, affects Y" (p. 631). It is important to note that modern mediation theory does not assume a significant correlation between X and Y [137]. All relationships were modeled as linear except for the one between arousal and pleasantness, for which a quadratic relationship was chosen. The results of these regression analyses, conducted for all participants, as well as for males and females separately, are shown in Table 2. In a first step, we calculated a mediator model and specified familiarity as a covariate. This model significantly explained around 26% of the variance of arousal (M) (p < .001) when both groups of participants were considered in the analysis. As expected, and in line with Berlyne's theory, complexity was positively correlated with arousal (a = .26, p = .008). The dependent mediator model was calculated in a second step and explained 55% of the variance in pleasantness (p < .001). There was no significant direct effect of complexity on pleasantness (c' = .10, p = .458), which represents how much a unit change in complexity affects pleasantness independent of its effect on arousal. With respect to the effect of arousal on pleasantness, the coefficient of the linear term was positive and significant (b_1 = 2.95, p = .009), whereas the coefficient of the quadratic term was negative and significant (b_2 = -.43, p = .002). Following the suggestion of Hayes and Preacher [132], the sample mean and one standard deviation above and below the mean were used to estimate the conditional indirect effect, that is, the significance of the indirect effect from complexity to pleasantness through arousal, conditional on specific values of complexity. A confidence interval was computed applying a bias-corrected resampling bootstrap technique with 5000 resamples. These values were negative at a 95% confidence level and thus statistically different from zero, demonstrating a linear instantaneous indirect effect: At all levels of picture complexity, an increase of complexity led to a decrease in pleasantness through the effect of complexity on arousal. A similar pattern of results was observed for males and females and when models without considering familiarity as a covariate were tested. Bearing in mind that causal path models should be treated with caution, especially in cross-sectional designs, an alternative reversed mediation model using pleasantness as the independent variable and complexity as the dependent variable was calculated. This model could only explain 17% of the variance of complexity (p = .001), which makes the pleasantness-arousal-complexity causal system a less likely candidate. Taken together, the current data suggests that arousal plays a significant role in the complexity-pleasantness relationship, supporting Berlyne's theory [7,10]. Moreover, complexity and pleasantness were linearly related when complexity was manipulated by the number of elements in an environmental scene, which is generally in line with the findings by Nadal et al. [16], although the current linear relationship was negative and not positive.

In order to investigate whether subjective ratings of complexity were associated with objective complexity as measured by compressed file size, a series of Spearman's rank-order correlations was conducted. Table 3 shows that all four measures of compressed file size were positively associated with subjective complexity. The correlations between the subjective ratings and TIFF file size (r_s = .53) and JPEG file size (r_s = .52) were the strongest, followed by the one of PNG file size (r_s = .46), whereas the one with the GIF file size was the lowest (r_s = .29) among the different types of

Table 2. Indirect effect of complexity (X) on pleasantness (Y) through arousal (M) calculated for IAPS pictures (N_{All} = 94, n_{males} = 92, $n_{females}$ = 92), modeled with a quadratic relation between M and Y and familiarity as a covariate.

		Model predicting arousal (M)				
		Coeff		*SE*		
Constant	f	3.64***		.58		
	m	4.58***		.50		
	All	4.27***		.53		
Familiarity	f	-.11		.08		
	m	-.39***		.07		
	All	-.29***		.07		
Complexity (X)	f	.31**		.11		
	m	.20*		.09		
	All	.26***		.10		
Summary of model predicting M	f	R^2 = .13***				
	m	R^2 = .35***				
	All	R^2 = .26***				
		Model predicting pleasantness (Y)				
Constant	f	-5.10*		2.52		
	m	-1.51		2.03		
	All	-3.57		2.20		
Familiarity	f	.73***		.10		
	m	.36**		.11		
	All	.57***		.11		
Complexity (X)	f	.10		.15		
	m	.12		.12		
	All	.10		.12		
Arousal (M)	f	3.14*		1.23		
	m	2.56*		1.06		
	All	2.95**		1.11		
Arousal squared (M*M)	f	-.42**		.14		
	m	-.43**		.14		
	All	-.43**		.14		
Summary of model predicting Y	f	R^2 = .53***				
	m	R^2 = .52***				
	All	R^2 = .55***				
		M_x	Θ_x	CI 95%		
	Pictures of low complexity	f	2.77	-.09	-.22	-.01
	m	2.77	-.09	-.22	-.02	
	All	2.80	-.09	-.22	-.02	
Pictures of moderate complexity	f	3.56	-.15	-.33	-.04	
	m	3.62	-.12	-.28	-.03	
	All	3.61	-.13	-.30	-.04	
Pictures of high complexity	f	4.35	-.21	-.50	-.05	
	m	4.48	-.15	-.36	-.03	
	All	4.42	-.18	-.42	-.04	

Table 2 (continued).

Note. X = independent variable; M = mediator variable, Y = dependent variable; *Coeff* = unstandardized coefficient; f = females; m = males; Θ_x = instantaneous indirect effect of X on Y through M at a specific value X = x; CI = confidence interval; $p < .05$, $p < .01$, $p < .001$.

compression formats. In general, the correlations between the ZIP and 7Z versions of the respective compression formats and subjective complexity yielded very similar results and are thus not reported here. One exception concerns the ZIP version of the JPEG format, which did not correlate significantly with subjective complexity. There were no significant correlations between the different compression formats and any other type of subjective rating, and further, no indications of any significant gender differences were found. In summary, the current results suggest that the correlation between subjective complexity and the original JPEG file size is among the strongest and that further transformations may not yield better results.

The observed positive correlations between compressed file size and subjective complexity are in line with earlier findings [101,102,104]. Considering the type of pictures used in the current experiment, i.e. environmental scenes, a comparison with the studies by Cavalcante et al. [104] and Forsythe et al. [74] seems to be appropriate. Cavalcante et al. [104] investigated the relationship between JPEG file size and subjective complexity for 74 days- and nightscapes of cities. Their findings indicated a slightly weaker positive relationship (r = .36) as the one observed in the current experiment (r_s = .52). Forsythe and colleagues [74] examined the relationship between two types of compression file size (JPEG and GIF) with subjective complexity ratings in response to natural pictures (N = 200) presented for 5 s. In their study, JPEG file size correlated somewhat more strongly with subjective complexity (r_s = .60) than in the current experiment. Forsythe et al. [74] also reported a similar correlation between GIF file size and natural pictures (r_s = .55), which could not be replicated with the current set of stimuli (r_s = .29).

Table 3 shows the results of Spearman's rank-order correlations between different types of edge detection measures and subjective ratings of familiarity, complexity, pleasantness and arousal (see Table S2 for inter-correlations between a representative set of objective complexity measures). Only the raw measures of the perimeter detection measure and Canny algorithm are depicted in Table 3 because the measures based on compressed file size yielded very similar results. The four measures of perimeter detection correlated stronger (all r_s ~ .44) with subjective complexity compared to the measures based on the Canny algorithm (all r_s ~ .30). Therefore, the current results support the findings by Forsythe et al. [74], who also reported a moderate correlation between perimeter detection and subjective complexity (r_s = .54) for a set of environmental scene photographs.

It further needs to be noted that we did not observe any significant positive correlations between participants' self-reported familiarity with the pictures and the four measures of perimeter detection. This result is in line with Forsythe et al. [102], who reported that perimeter detection and its relation to

Table 3. Spearman's rank-order correlations between a set of 13 measures of objective complexity applied to IAPS pictures (N = 96) and ratings of familiarity, complexity, pleasantness and arousal.

Measure		JPEG	GIF	PNG	TIFF	PERI-RAW	CANNY-RAW	ENTROPY	RMS-CONTR *M*	RMS-CONTR *SD*	RMS-CONTR α	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG β
Familiarity	f	.03	.08	.06	.08	.18	.15	.13	.08	.07	.07	.03	.00	.03
	m	.02	.08	.03	.04	.11	.11	.14	.06	.05	.05	.002	-.05	-.01
	All	.03	.07	.04	.05	.16	.14	.12	.08	.07	.07	.04	-.01	.03
Complexity	f	.48*	.28	.44*	.51*	.41*	.27	.30	.54*	.38*	.53*	.32*	.30*	.33*
	m	.52*	.30	.43*	.49*	.44*	.27	.26	.58*	.45*	.57*	.31*	.34*	.34*
	All	.52*	.29	.46*	.53*	.45*	.29	.28	.59*	.43*	.57*	.34*	.34*	.36*
Pleasantness	f	.08	.09	.08	.11	.07	.14	-.002	.07	.18	.14	.07	.11	.08
	m	.10	.08	.10	.12	.10	.15	.004	.10	.21	.16	.06	.11	.07
	All	.09	.09	.09	.11	.09	.15	.005	.08	.20	.16	.07	.11	.08
Arousal	f	.06	.04	.02	.03	.06	.03	.16	.11	-.03	.05	-.07	-.12	-.06
	m	.13	.05	.12	.13	.05	-.05	.14	.20	.04	.12	-.01	-.01	.02
	All	.12	.06	.11	.11	.08	-.01	.14	.18	.02	.10	-.03	-.04	.02

Note. * $p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; PERI = perimeter detection; RMS = root mean square; CONTR = contrast; CONG = congruency; all *dfs* ≥ 84 for females; all *dfs* ≥ 83 for males; all *dfs* ≥ 84 for all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

subjective complexity remain unaffected by the self-reported familiarity with the content of a picture. Furthermore, none of the compression file sizes and edge detection measures of complexity correlated significantly with pleasantness, which does not support the findings by Forsythe et al. [74], who reported an association between GIF compression and beauty ratings of various types of visual stimuli.

Finally, we found several positive associations between subjective complexity and objective measures of complexity, which either have not been explored yet or have only rarely been used before. The entropy of the grayscale pictures correlated weakly with subjective complexity (r_s = .28), which is comparable to the correlation we detected for the raw measure of the Canny edge detection algorithm. Similar correlations were observed for the three measures of edge detection based on phase congruency (all r_s ~ .34). Among all applied objective complexity measures related to edge detection, subjective complexity correlated best with the mean contrast values of the RMS contrast map (r_s = .59) and the α-measure of the RMS contrast values (r_s = .57). These results clearly replicate those of Cavalcante et al. [104], who reported correlations of r = ~ .60 between the three RMS contrast measures and subjective complexity ratings of streetscape images, revealing the potential future use of RMS contrast measures as reliable indicators of subjective complexity. It is worth noting that we also found weak, non-significant, indications for a positive association between pleasantness and the standard deviation of the mean RMS contrast values (all r_s ~ .20). Nevertheless, the current data suggests that correlations between objective measures of complexity and subjective ratings other than complexity are weak, and consequently, that the applicability of objectives measures of complexity to the prediction of subjective pleasantness ratings appears limited for photographs of environmental scenes.

Linear multiple regression analyses were conducted to further investigate the inter-relationships between subjective ratings and objective measures of complexity. Table 4 summarizes the results of a stepwise linear regression model with subjective complexity as the criterion (dependent variable) and a set of ten variables as predictors, including familiarity, pleasantness and arousal as well as a set of seven measures of objective complexity. Familiarity was entered in the first step. Due to issues of multicollinearity, the following six measures of objective complexity were excluded: TIFF and PNG file size, *SD* and α-measures of RMS contrast, and *SD* and β-measures of edge detection based on phase congruency. This regression model allowed addressing the question of whether objective measures can predict subjective complexity better than, for example, subjective arousal. The adjusted R^2 value indicated that the model predicted 51% of the variance, with the mean values of the RMS contrast measure as the strongest predictor (β = .55, p < .001), followed by familiarity (β = -.48, p = .001). The RMS contrast measure accounted for around 30% of the variation in subjective complexity. A similar model considering only objective measures of complexity as predictors (Table S3) revealed that the mean of the RMS contrast values was the only significant predictor of subjective complexity (β = .54, p < .001), explaining approximately 28% of the variance. Similar regression analyses conducted for males and females separately revealed analogous results except that arousal was a third significant predictor in the model for females (β = .23, p = .012). Taken together, these results support the view that familiarity is a significant predictor of subjective complexity that cannot be ignored in any complex models, and further, that the mean RMS contrast values are the strongest predictor of subjective complexity among the current set of objective complexity measures applied to affective environmental scenes.

Table 4. Summary of linear stepwise regression analysis for ten variables predicting subjective complexity ratings of IAPS pictures (N = 70).

Variable	*B*	*SE B*	B
Step 1			
Constant	5.15	.35	
Familiarity	-.35	.08	-.47***
Adjusted R^2	.21		
F	18.98***		
Step 2			
Constant	4.02	.33	
RMS contrast *M*	20.66	3.17	.55***
Familiarity	-.37	.06	-.48***
Adjusted R^2	.51		
F	36.58***		
ΔR^2	.30		

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; *B* = unstandardized regression coefficient; SE = standard error; β = standardized regression coefficient; ΔR^2 = difference in the proportion of variance explained.

Experiment 2

Methods

Participants. Forty German-speaking students (20 males, 24.4 ± 4.1 years; age range 19-38 years; 20 females, 22.0 ± 1.5 years, age range 20-26 years) participated in the study in return for course credit. Most of these students were enrolled in a psychology degree and none of them was pursuing an art history degree or was an arts expert. Participants did not take part in any other experiment in the current study, and they all had normal or corrected-to-normal visual acuity.

Materials. Ninety-six high-quality digital reproductions of colored oil and acryl paintings were downloaded from two digital image libraries (*prometheus*, http://prometheus-bildarchiv.de, and *ARTStor*, http://www.artstor.org/index.shtml). These representational paintings comprised different styles from the end of the 18th to the beginning of the 20th century, including renowned artists such as Achenbach, Cezanne, Courbet, Friedrich, Gauguin, Gericault, Goya, Manet, Monet, Turner, and Van Gogh (see Supporting Information Stimulus List S2). The semantic content of the paintings was similar to the one of the IAPS pictures used in Experiment 1, depicting human beings in different situations of everyday life (excluding erotic scenes), landscapes, animals and plants. Paintings that were highly familiar and for which the semantic content could not be easily understood, as judged by the co-authors, were not included in the stimulus set. As in Experiment 1, the paintings were pre-selected, as judged by the co-authors, to fall into one of the four quadrants spanned by arousal and pleasantness. For this purpose, the range of the semantic content was matched as much as possible with the one of environmental scenes. In order to achieve sufficient variation in subjective arousal, paintings with dull colors where preferably pre-selected for the low-arousing emotion quadrants. Half of the paintings in each hedonic category (n = 24) depicted a figure-ground composition and half a complex scene. The paintings were all in landscape format and saved in the same size (1024 x 768 pixels) and resolution (72dpi) in JPEG format (maximum quality) as the IAPS pictures used in Experiment 1. Due to the natural preponderance of figure-ground compositions in portrait format rather than in landscape format, especially when paintings contained human faces, several paintings were cut to landscape format using Adobe Photoshop CS5 software. Other modifications concerned the presence of signatures or frames, which were removed. It was made sure that all modifications yielded natural representations of paintings without any deformations. Since we aimed to preserve the emotional contents of the paintings, we decided not to control for variations of brightness within the stimulus set. The paintings were not modified in any other way.

In order to rule out that any observed differences between the results regarding paintings and IAPS pictures were due to differences in brightness, an arithmetic mean model was applied to calculate the average brightness of the 96 paintings and 96 IAPS pictures in the RGB (red-green-blue) color space in Matlab. Specifically, brightness can be thought of as the arithmetic mean μ of the red, green and blue color coordinates (μ = (R+G+B)/3). The sets of paintings (M = .38, SD = .14) and IAPS pictures (M = .37, SD = .14) did not significantly differ in their average RGB pixel levels, $t(190) = -.35$, $p = .726$.

Two self-developed questionnaires were used, one asking specific questions referring to the experiment and one probing general interest in arts and expertise in visual arts. The former questionnaire comprised two questions about whether participants recognized the styles of the paintings and could name the century of their creation. They were also asked to indicate the degree of their general liking for the paintings on a seven-point Likert scale (from "not at all" to "very much"). Moreover, participants reported the general level of difficulty associated with giving ratings of complexity on a seven-point Likert scale (from "very easy" to "very difficult"). Finally, participants estimated the percentage of paintings they had seen at least once before the experiment (ranging from 0% to 100%). The self-developed questionnaire on interest in arts and expertise in visual arts comprised three parts. A set of eleven questions was used to examine participants' interest in visual arts (nine-point Likert scale, from "fully disagree" to "fully agree"), followed by a part in which participants had to indicate whether they are familiar with the names of a range of artists as well as to name their nationality and their associated style by free verbal responses. In the final part, participants were confronted with representations of six paintings and again had to indicate their familiarity with the paintings and to name the artist and the style.

Procedure. In order to be able to compare results across the two different visual stimuli sets used in this study, the procedure of Experiment 2 was generally similar to the one in Experiment 1, with the exception that the instructions for familiarity ratings were as follows: "Please rate your familiarity with the painting." In addition, after participants had finished the experiment, they filled a self-developed one-page questionnaire relating directly to the experiment followed by a four-page questionnaire assessing general interest in arts and

expertise in visual arts. These additional questionnaires extended the length of the experimental session to 90 minutes. As in Experiment 1, the original paintings in JPEG format were transformed into a set of twelve different compression formats of which the file sizes were determined. Several edge detection measures and the entropy were calculated by using the same procedures as described in Experiment 1.

Results and Discussion

With regard to the questionnaire data on participants' liking for the paintings used in the experiment and the level of difficulty of giving complexity judgments, all except one out of four Shapiro-Wilk normality tests revealed that the data was not normally distributed (ps < .01). Further analyses showed that males (n = 20) and females (n = 20) did not differ in their general liking for the paintings, Mann-Whitney U = 181.5, p = .608, r = -.08, and also not in their self-reported degree of difficulty in judging the complexity of the paintings, Mann-Whitney U = 155.5, p = .214, r = -.20. Note that r refers to an effect size estimate that is derived by converting test statistics into z-scores and dividing by the square root of the number of total observations, $r = z/\sqrt{N}$ ([138], p. 19). The standard values of abs(r) for small, medium and large effect sizes are as follows: small: r = .10; medium: r = .30; large: r = .50.

Participants reported to have seen approximately 5-10% of the paintings prior to the experiment on average. Regarding the results of the questionnaire on the self-reported interest in visual arts, we did not observe a significant difference between males (n = 19, M = 6.24, SD = 1.5) and females (n = 20, M = 5.4, SD = 1.34), $t(37)$ = 1.85, p = .072, r = .29. One male participant did not respond to all the items of the art interest questionnaire. Screening of the other parts of the questionnaire (art expertise) indicated that none of the participants was an art expert.

As in Experiment 1, subjective ratings were averaged across participants for each painting prior to the main analysis. In order to test whether participants rated each stimulus in a consistent fashion, inter-rater reliability was assessed by computing the average measure intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency [133,134]. We observed high inter-rater reliability for all four rating scales, as evidenced by the following results based on all participants: familiarity (ICC(2, k) = .86, 95% CI [.82, .90]), complexity (ICC(2, k) = .93, 95% CI [.90, .96]), pleasantness (ICC(2, k) = .96, 95% CI [.95, .97]) and arousal (ICC(2, k) = .90, 95% CI [.87, .92]). The following results were obtained for females: familiarity (ICC(2, k) = .82, 95% CI [.76, .87]), complexity (ICC(2, k) = .87, 95% CI [.83, .90]), pleasantness (ICC(2, k) = .94, 95% CI [.92, .95]) and arousal (ICC(2, k) = .84, 95% CI [.79, .88]). In the group of males, inter-rater reliability for the four scales was as follows: familiarity (ICC(2, k) = .71, 95% CI [.61, .79]), complexity (ICC(2, k) = .89, 95% CI [.86, .92]), pleasantness (ICC(2, k) = .92, 95% CI [.90, .95]) and arousal (ICC(2, k) = .80, 95% CI [.73, .85]).

Data screening of the variables of objective complexity for outliers revealed between zero and eleven outliers per variable. Outliers were defined by values ± 2 SD from the mean and removed. A series of Shapiro-Wilk normality tests indicated that four variables deviated significantly from normality after removal of outliers (all ps < .03). The majority of the variables of self-reported familiarity, complexity, pleasantness and arousal did not deviate significantly from a normal distribution when all participants were considered together (all ps > .04), or when males (all ps > .09) and females (all ps > .01) were considered separately.

The distribution of the paintings in the two-dimensional emotion space (Figure 2a) was very similar to the one of IAPS pictures observed in Experiment 1 (Figure 1a), showing a quadratic relationship between arousal and pleasantness, R^2 = .30, $F(2,88)$ = 18.88, p < .001, $y = -2.42x + .28x^2 + 8.69$. However, the strength of the relationship was weaker in the set of paintings than the one observed for IAPS pictures. A further comparison between the two emotion spaces indicated that the range of arousal for paintings (min. M = 2.50, max. M = 4.85, range = 2.35) was smaller than the one seen for IAPS pictures (min. M = 2.53, max. M = 5.75, range = 3.22). The range of pleasantness associated with the paintings (min. M = 2.13, max. M = 5.55, range = 3.42) was also smaller than the one associated with IAPS pictures (min. M = 1.39, max. M = 6.08, range = 4.69). Fewer paintings were regarded as highly unpleasant and high-arousing. Nevertheless, the current data suggest the presence of a negativity bias [135] of emotional processing of representational paintings. This finding may be explained by the fact that the environmental scenes used in Experiment 1 had similar semantic contents as the current representational paintings and that both types of stimuli presumably induced emotional processes relevant to survival in nature.

Table 5 and Figure 2 show that, when males and females were considered together, the nature of the relationships between the subjective ratings of familiarity, complexity, pleasantness and arousal was generally similar to the one observed in Experiment 1. However, in contrast to Experiment 1, familiarity was not significantly associated with complexity (r_s = -.07), but the data revealed a significant negative association with arousal (r_s = -.28) and a significant positive association with pleasantness (r_s = .78). The latter correlations were of a roughly similar magnitude as the ones observed in the set of IAPS pictures. Moreover, complexity was significantly positively related with arousal (r_s = .54). The non-significant negative relationship between complexity and pleasantness was weaker (r_s = -.10) than in Experiment 1 (r_s = -.18). In order to control for effects of familiarity, partial Spearman's rank-order correlations (Table S4) were performed on the subjective reports of complexity, pleasantness and arousal. A comparison between results presented in Tables 5 and S4 revealed that the positive association between complexity and arousal gained slightly in strength (r_s = .59), whereas the relationship between complexity and pleasantness did not change.

In light of Berlyne's theory [9], the finding of a positive association between complexity and arousal could be corroborated in both genders (r_s = .59), as was the case for IAPS pictures in Experiment 1 (r_s = .27). This association between complexity and arousal was significantly stronger for paintings (z = -2.63, p = .009) when effects of familiarity were

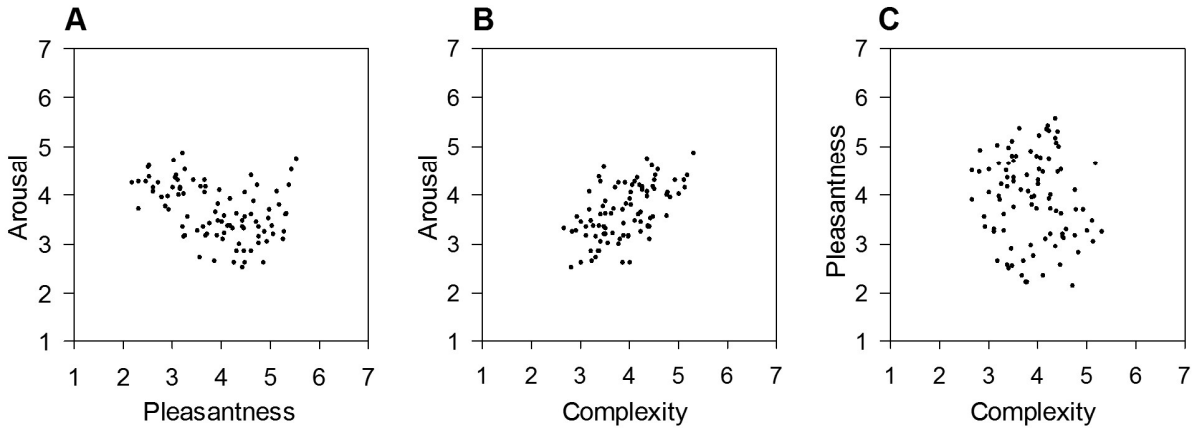

Figure 2. Relationships between pleasantness, arousal and complexity in a set of representational paintings. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Distribution of paintings in the pleasantness-arousal space based on mean ratings per picture. B) Relationship between mean complexity and arousal ratings. C) Relationship between mean complexity and pleasantness ratings.

doi: 10.1371/journal.pone.0072412.g002

Table 5. Spearman's rank-order correlations between ratings of familiarity, complexity, pleasantness and arousal in response to representational paintings ($N_{All} \geq 89$, $n_{males} \geq 89$, $n_{females} \geq 89$).

Measure		Familiarity	Complexity
Complexity	f	-.33*	
	m	.16	
	All	-.07	
Pleasantness	f	.81*	-.26*
	m	.66*	.10
	All	.78*	-.10
Arousal	f	-.45*	.56*
	m	-.10	.43*
	All	-.28*	.54*

Note. * $p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

partialed out. Furthermore, we did not observe a clear linear relationship between complexity and pleasantness when males and females were analyzed separately (Figure S2) and when familiarity effects were controlled for, which is in line with results of Experiment 1. A significant negative linear relationship was only present in the group of females when familiarity was not considered in the analysis. This finding does not lend support to Nadal et al.'s [16] theory of a positive linear relation between complexity and hedonic value if complexity is manipulated by the number and variety of elements. One may hypothesize that this discrepancy in research findings could be related to the fact that different hedonic measures were applied in the two studies (beauty vs. pleasantness). Future studies may thus investigate whether the three relationships between different dimensions of complexity and beauty (linear, U-shaped and inverted U-shaped) described by Nadal et al. [16] hold true for other types of hedonic measures besides beauty. Another alternative explanation for the current finding may simply be attributed to the different degree of aesthetic quality and motivational relevance of the visual materials under investigation compared to those used by Nadal et al. [16], who avoided including stimuli inducing strong affect.

The current data also revealed several gender differences with regard to subjective ratings (Figure S2), which were not present in Experiment 1. In general, familiarity correlated stronger with complexity, pleasantness and arousal in females compared to males. The relationship between familiarity and complexity was reversed in the two groups, $z = -3.31$, $p < .001$. In females, the relationship between familiarity and complexity was negative ($r_s = -.33$), while it was positive in males ($r_s = .16$). Familiarity and pleasantness were also more strongly associated in females ($r_s = .81$) than in males ($r_s = .66$), $z = 2.19$, $p = .029$. Moreover, the negative relationship between familiarity and arousal was stronger in females ($r_s = -.45$) than in males ($r_s = -.10$), $z = -2.52$, $p = .012$. Last, the direction and strength of the relationship between complexity and pleasantness (Figure S2) differed when males ($r_s = .10$) and females ($r_s = -.26$) were analyzed separately, $z = -2.4$, $p = .016$. The current findings thus corroborate earlier reports of gender differences in visual art appreciation [86,139,140], highlighting the necessity to consider gender in the study of empirical aesthetics.

As in Experiment 1, a mediation model was calculated using the macro "MEDCURVE" in SPSS [132] in order to examine whether arousal mediates the relationship between complexity and pleasantness. The relationships between complexity and arousal as well as between complexity and pleasantness were modeled as linear, while the relationship between arousal and pleasantness was modeled as quadratic. Results of these analyses, separately for males, females and all participants,

are presented in Table 6. In a first step, we calculated a mediator model and specified familiarity as a covariate. This model significantly explained 44% of the variance in arousal (M) (p < .001) when both groups of participants were considered in the analysis. As expected, and in line with Berlyne's theory and results of Experiment 1, complexity was positively correlated with arousal (a = .49, p < .001). In a second step, the dependent variable model was calculated and explained 59% of the variance in pleasantness (p < .001). The direct effect of complexity on pleasantness was not significant (c' = .11, p = .346), which we also observed in Experiment 1. With respect to the effect of arousal on pleasantness, the coefficient of the linear term and the coefficient of the quadratic term were not significant. Only the effect of familiarity was highly significant (p < .001). There were no indications of any significant instantaneous indirect effects when males and females were analyzed together, which stands in clear contrast to results reported in Experiment 1. However, a further mediation model conducted without considering familiarity as a covariate revealed significant linear indirect effects of complexity through arousal on pleasantness based on data of all participants. As in Experiment 1, we tested an alternative mediation model (pleasantness-arousal-complexity) which could only explain around 36% of the variance, implying that the complexity-arousal-pleasantness path model may better capture the underlying perceptual and cognitive processes.

By using a bootstrapping approach, instantaneous indirect effects of complexity on pleasantness through arousal were observed when males and females were analyzed separately. Interestingly, this mediation effect was only present for pictures of moderate and high complexity in both groups, suggesting that for pictures of moderate and high complexity an increase in complexity led to a decrease in pleasantness through an increase in arousal. For pictures of low complexity, no such mediation effect was found. Further, it needs to be noted that the direct effect of complexity on pleasantness was nearly significant in females (c' = .22, p = .074). We believe that a gender-wise analysis is clearly better suited to detect mediation effects since we have shown that, for example, correlations between familiarity and the respective subjective ratings differed between males and females in the current experiment. Important information would have been lost if only results based on the pooled data had been presented.

A critical interpretation of the current results suggests that, on the one hand, the underlying mechanisms of the perception of visual artistic stimuli may differ from the perception of environmental scenes. For example, results of Experiment 1 provide a very coherent picture with regard to significant indirect effects of complexity through arousal on pleasantness as well as with regard to the absence of a direct effect of complexity on pleasantness. In the current experiment involving paintings, significant indirect effects only emerged when males and females were analyzed separately, and further, weak indications of a direct effect were observed in females. Moreover, the relationship between complexity and arousal was stronger in the perception of representational paintings compared to the perception of environmental scenes. On the other hand, albeit the noted differences between the

Table 6. Indirect effect of complexity (X) on pleasantness (Y) through arousal (M) calculated for representational paintings (N_{All} = 86, n_{males} = 87, $n_{females}$ = 88), modeled with a quadratic relation between M and Y and familiarity as a covariate.

		Model predicting arousal (M)				
		Coeff		*SE*		
Constant	f	2.88***		.54		
	m	2.74***		.38		
	All	2.80***		.42		
Familiarity	f	-.34**		.10		
	m	-.34**		.10		
	All	-.38***		.10		
Complexity (X)	f	.48***		.09		
	m	.44***		.07		
	All	.49***		.07		
Summary of model predicting M	f	R^2 = .40***				
	m	R^2 = .33***				
	All	R^2 = .44***				
		Model predicting pleasantness (Y)				
Constant	f	-2.75		2.03		
	m	-.33		2.36		
	All	-1.20		2.38		
Familiarity	f	1.24***		.12		
	m	.87***		.13		
	All	1.21***		.14		
Complexity (X)	f	.22		.12		
	m	.15		.11		
	All	.11		.12		
Arousal (M)	f	1.59		1.02		
	m	1.15		1.38		
	All	1.03		1.28		
Arousal squared (M*M)	f	-.26*		.13		
	m	-.22		.20		
	All	-.19		.18		
Summary of model predicting Y	f	R^2 = .69***				
	m	R^2 = .51***				
	All	R^2 = .59***				
		M_x	Θ_x	CI 95%		
	Pictures of low complexity	f	3.47	-.13	-.30	.03
	m	3.13	-.11	-.25	.05	
	All	3.30	-.10	-.28	.08	
Pictures of moderate complexity	f	4.08	-.20	-.37	-.08	
	m	3.80	-.17	-.32	-.05	
	All	3.93	-.16	-.34	.02	
Pictures of high complexity	f	4.69	-.28	-.49	-.12	
	m	4.47	-.23	-.46	-.06	
	All	4.55	-.22	-.50	.05	

Table 6 (continued).

Note. X = independent variable; M = mediator variable; Y = dependent variable; *Coeff* = unstandardized coefficient; f = females; m = males; All = all participants; Θ_x = instantaneous indirect effect of X on Y through M at a specific value X = x; CI = confidence interval; $p < .05$, $p < .01$, $p < .001$.

results of Experiments 1 and 2, which are presumably partly due to gender effects with regard to familiarity ratings of paintings, the presence of significant indirect effects when viewing environmental scenes and representational paintings lends support to Berlyne's theory [7,10]. Future studies may investigate the presence of a mediation effect with different types of visual stimuli and hedonic measures to see whether the current findings can be generalized.

The relationships between subjective and objective complexity as measured by compressed file size were investigated by Spearman's rank-order correlations. Table 7 shows that none of these correlations was significant (i.e., JPEG, GIF, PNG and TIFF formats as well as their 7z and ZIP versions which are not reported here), which stands in clear contrast to the results reported in Experiment 1. Thus, these findings are not in line with Forsythe et al. [74], who reported that JPEG file size (r_s = .40) and GIF file size (r_s = .47) correlated significantly with subjective complexity judgments of representational paintings (N = 148). Furthermore, the current twelve compression file sizes neither correlated with familiarity, pleasantness nor arousal, which is in agreement with the results in relation to IAPS pictures investigated in Experiment 1 (Table 3). The lack of a significant relationship between compressed file size and subjective complexity could be due to the fact that the current set of paintings was pre-selected to vary considerably in the number of elements present in a visual scene. A figure-ground composition, such as a portrait, may be judged as relatively simple but even a relatively uniform background may be difficult to compress because of the natural variability of the brush marks present in most of these paintings. Consequently, the application of compression file size as a measure of objective complexity may depend on the specific type of complexity dimension under investigation as well as on the stimulus type. Forsythe et al. [74] did not pre-select their stimuli according to one or more complexity dimensions (M. Nadal-Roberts, personal communication, October 15, 2012), thus making it difficult to establish a direct comparison with their results.

Table 7 also presents the Spearman's rank-order correlations between the four subjective variables and the set of edge detection measures as well as entropy. Neither the raw nor the compressed file size measures of perimeter detection yielded any significant relationships with subjective complexity, which can partly be explained by the nature of the paintings under investigation. For example, perimeter detection may be sensitive to coarse features of the background in a figure-ground composition and thus weaken the positive association with subjective complexity observed in Experiment 1. Such an interpretation of the results would explain why the current findings differ from those reported by Forsythe et al. [74], who found a positive association between perimeter detection and subjective complexity (r_s = .37). Bearing in mind that Forsythe et al. [74] did not specifically include figure-ground compositions, it may be that perimeter detection was still successful in predicting subjective complexity to a moderate extent.

The measures related to the Canny algorithm correlated negatively (but not significantly) with subjective complexity (r_s = -.23), a result which clearly differs from the positive association between these measures as presented in Experiment 1 (r_s ~ .29). A closer inspection of the paintings after analysis by the Canny algorithm revealed that fine features of the background, due to the individual brush strokes, were indeed largely detected, which may explain the negative association with subjective complexity. Moreover, there was no significant association between the entropy of a grayscale image and subjective complexity. Interestingly, the mean RMS contrast values correlated only weakly with subjective complexity, but positive correlations were observed for the standard deviation of the RMS contrast values (r_s = .24) and the RMS contrast α-measure (r_s = .25). However, the degree of these correlations was considerably smaller than the one found for IAPS pictures in Experiment 1 (r_s ~ .60). The standard deviation of the mean values of edge detection based on phase congruency yielded the strongest (and only significant) correlation (r_s ~ .38) among all measures. It seems that objective measures of complexity including a measure of dispersion, such as the *SD* measures of RMS contrast and edge detection based on phase congruency as well as the related α- and β-measures, worked better for the current set of artworks than for environmental scenes. In summary, our findings indicate that objective measures which were successful in predicting subjective complexity of environmental scenes were much less efficient for representational paintings. Inter-correlations between a representative set of objective measures of complexity can be found in the Supporting Information (Table S5).

The present results not only revealed significant correlations between measures of subjective and objective complexity, but also indications of relationships between objective measures of complexity and other types of subjective ratings. The correlation between the standard deviation RMS contrast measure and arousal was significant (r_s = .37), which stands in clear contrast to the findings of Experiment 1. In addition, a negative relationship between Canny edge detection measures and subjectively reported arousal (all r_s ~ -.23) was observed. The finding that edge detection measures correlated with arousal in paintings but not in IAPS pictures may be explicable by the fact that the correlation between subjective arousal and complexity was stronger in paintings (r_s = .54) than in IAPS pictures (r_s = .36). Furthermore, the data showed indications of weak correlations between subjective pleasantness and measures of Canny edge detection (r_s =.26) and the standard deviation of the RMS contrast values (r_s = -.28), respectively. This lends some support to the finding by Forsythe et al. [74], who reported a significant correlation between objective complexity as measured by compressed file size and subjective ratings of beauty. Last but not least, indications of a positive correlation between familiarity and JPEG, PNG and

Table 7. Spearman's rank-order correlations between a set of 13 measures of objective complexity applied to representational paintings (N = 96) and ratings of familiarity, complexity, pleasantness and arousal.

Measure		JPEG	GIF	PNG	TIFF	PERI	RAW	CANNY	RAW	ENTROPY	RMS-CONTR *M*	RMS-CONTR *SD*	RMS-CONTR α	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG β
Familiarity	f	.05	-.07	.07	.08	.02	.20	.10	-.07		-.16	-.14	-.08	-.06	-.05	
	m	.30	.05	.18	.24	-.03	.01	.10	.05		.04	.04	-.03	.04	.01	
	All	.22	.05	.21	.23	.06	.13	.12	.05		-.03	.001	-.05	-.01	-.02	
Complexity	f	-.02	.05	-.09	-.08	.11	-.28	.03	.09		.25	.23	.27	.37*	.30	
	m	.02	.01	-.04	.002	.14	-.22	.09	.09		.20	.22	.24	.37*	.29	
	All	.02	.06	-.06	-.02	.15	-.23	.10	.11		.24	.25	.25	.38*	.30	
Pleasantness	f	.20	-.07	.11	.18	.03	.30	.03	-.06		-.27	-.17	-.09	-.10	-.09	
	m	.16	-.14	.07	.12	-.04	.16	.01	-.13		-.24	-.18	-.05	-.05	-.03	
	All	.17	-.09	.10	.15	.03	.26	.05	-.11		-.28	-.19	-.07	-.08	-.06	
Arousal	f	-.09	.02	-.13	-.12	.02	-.23	-.02	.12		.34	.23	.23	.21	.26	
	m	-.07	.03	-.10	-.11	-.02	-.25	.07	.09		.31	.23	.13	.16	.14	
	All	-.10	.05	-.09	-.12	.03	-.23	.04	.14		.37*	.27	.23	.24	.25	

Note. *$p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; PERI = perimeter detection; RMS = root mean square; CONTR = contrast; CONG = congruency; all *dfs* ≥ 81 for females; all *dfs* ≥ 79 for males; all *dfs* ≥ 80 for all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

TIFF file sizes were observed in males, indicating that objective measures of complexity are not completely unaffected by familiarity as previously reported by Forsythe et al. [102]. In summary, the pattern of results observed for the current set of representational paintings differs in several respects from the results in relation to IAPS pictures (Experiment 1), drawing attention to the specificity of digital reproductions of representational art and environmental scenes within a common framework of complexity.

A stepwise regression analysis was conducted in order to examine whether objective measures of complexity are better predictors of complexity than, for example, subjective arousal. As in Experiment 1, we included seven measures of objective complexity (JPEG and GIF file size, perimeter detection, Canny algorithm, entropy, *SD* measures of RMS contrast and phase congruency) as well as arousal and pleasantness in the as predictors in the stepwise regression analysis. Familiarity was entered in a first step, although it was not significantly correlated with subjective complexity when all participants were considered together in the analysis. This procedure was chosen in order to guarantee comparability across the four experiments. The results, presented in Table 8, yielded a significant model (adjusted R^2 = .40) based on two significant predictors, namely arousal (β = .56, $p < .001$) and the standard deviation of the phase congruency edge detection measure (β = .28, p = .013). A comparison with results presented in Experiment 1 showed two main differences: First, subjective arousal was a better predictor of complexity ratings of representational paintings than the best objective measure, which was not the case for the set of IAPS pictures. Second, the mean RMS contrast values were the strongest predictor in Experiment 1, whereas the *SD* measure of edge detection based on phase congruency yielded the best results in the current experiment.

Similar stepwise regression analyses were conducted for both genders separately. In general, the results were comparable to those described above, with arousal and the *SD* measure of phase congruency as significant predictors of subjective complexity. However, the *SD* RMS contrast measure was identified as a significant suppressor variable in both multiple regressions. Following Pandey and Elliott [141], the

Table 8. Summary of linear stepwise regression analysis for ten variables predicting subjective complexity ratings of representational paintings (N = 59).

Variable	*B*	*SE B*	β
Step 1			
Constant	4.28	.56	
Familiarity	-.14	.20	-.10
Adjusted R^2	-.01		
F	.53		
Step 2			
Constant	.69	.79	
Arousal	.75	.14	.63***
Familiarity	.15	.17	.11
Adjusted R^2	.34		
F	15.76***		
ΔR^2	.35		
Step 3			
Constant	-.32	.85	
Arousal	.67	.13	.56***
SD phase congruency	22.37	8.74	.28*
Familiarity	.22	.16	.15
Adjusted R^2	.40		
F	13.74***		
ΔR^2	.07		

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; *B* = unstandardized regression coefficient; SE = standard error; β = standardized regression coefficient; ΔR^2 = difference in the proportion of variance explained.

current type of suppression could be identified as negative suppression, meaning that the *SD* RMS contrast measure was positively correlated with other predictors and the outcome variable but had a negative beta weight when entered into the regression model. For females, the model yielded an adjusted R^2 of .41, $F(4,57) = 11.42$, $p < .001$, with the following predictors: familiarity ($\beta = -.06$, $p = .61$), arousal ($\beta = .51$, $p < .001$), *SD* phase congruency ($\beta = .39$, $p = .001$) and *SD* RMS contrast ($\beta = -.25$, $p = .038$). The final model based on the data of males also explained 41% of the variance in subjective complexity, $F(4,57) = 11.44$, $p < .001$, including the following predictors: familiarity ($\beta = .26$, $p = .013$), arousal ($\beta = .53$, $p < .001$), *SD* phase congruency ($\beta = .43$, $p = .001$) and *SD* RMS contrast ($\beta = -.26$, $p = .035$).

As a final step, a multiple stepwise regression model only including objective measures of complexity was employed to predict subjective complexity ratings averaged over all participants. The results (Table S6) suggest that the *SD* measure of phase congruency ($\beta = .69$, $p < .001$) was the strongest predictor, followed by GIF compression file size ($\beta = .39$, $p = .003$) and the SD measure of RMS contrast ($\beta = -.30$, $p = .039$). Only 25% of the variance in subjective complexity could be explained, $F(3,60) = 7.90$, $p < .001$. Again, the *SD* measure of RMS contrast was identified as a negative suppressor variable. A comparison between the current results and those presented in Experiment 1 suggest that subjective complexity in response to IAPS pictures was easier to predict by means of objective measures because the model based on one predictor (mean RMS contrast values) was as successful as the current model based on three predictors. In both experiments, measures based on RMS contrast values contributed as significant predictors of visual complexity, highlighting their future use in studies involving objective complexity measures.

Experiment 3

Methods

Participants. Thirty-six German-speaking psychology students (18 males, 23.6 ± 2.8 years, age range 20-30 years; 18 females, 22.2 ± 3.9 years, age range 19-33 years) participated in the experiment. Participants received course credit in return for their participation and did not take part in any other experiment in this study. All participants were enculturated in Western tonal music from their birth onwards and non-musicians, i.e., they were neither playing a musical instrument at the time of the experiment nor had they played an instrument for longer than four years in their childhood or youth. Sixteen participants had played an instrument in their past (musical training = 1.10 ± 1.40 years). All participants reported normal hearing.

Materials. The ninety-two musical excerpts in WAV format were taken from the stimulus set described in Marin et al. [83]. The stimuli can be considered as representative of the 19th century Romantic piano solo music repertoire and comprised composers such as Brahms, Chopin, Grieg, Mendelssohn, Schubert and Schumann (see Supporting Information Stimulus List S3). Based on the distribution of the musical excerpts in the two-dimensional emotion space reported in Experiment 1 in Marin et al. [83], 23 stimuli were chosen to exhibit maximal variation in arousal and pleasantness. Excerpts thus included the outer representatives of each emotion quadrant spanned by arousal and pleasantness to obtain large differences in arousal and/or pleasantness and to cover the whole emotion space of this particular musical style. All stimuli were cut to 25 s of length (keeping the original beginning of the excerpts and removing parts of the ending of the excerpts selected from [83]), and a 500 ms fade in/out was added to the excerpts using Audacity 1.3.14 software. Importantly, in the pre-selection process only stimuli in which the emotional content was not largely changing over the 25 s were selected because the aim was to select stimuli that induce a relatively unambiguous emotion. The sounds were not modified in any other way.

A self-developed two-page questionnaire was used to explore participants' musical background. Participants answered questions referring to their amount of musical training and indicated how often they listen to a particular musical style ($n = 14$, ranging from classical, pop, Jazz to music of foreign cultures) on a seven-point Likert scale (from "never" to "very often"). In a similar way, they reported how often they listen to music actively (without doing other activities at the same time), passively (while doing other activities at the same time) or in a live-concert setting. For each of these ratings, participants named at least one musical style to which the rating mostly referred. Furthermore, they named their least and most preferred musical style and reported on the role music plays in their daily lives (seven-point Likert scale, with "no role… a big role" as anchors). Three questions referred directly to the experiment. Participants had to indicate whether they recognized the musical style and, if so, they were asked to name it. They also reported the degree of liking for the music on a seven-point Likert scale (from "not at all" to "very much") and the degree of difficulty of judging the complexity of the music (from "very easy" to "very difficult").

Procedure. In general, the design and procedure of the behavioral experiment were identical to Experiments 1 and 2 of the current study, except that participants were specifically screened for their amount of musical training during the recruitment process and further, that all participants were tested individually in a quiet room. Familiarity ratings referred to the specific musical excerpt and the instructions said: "Please rate your familiarity with the musical excerpt." Musical excerpts were blocked according to emotional contents as was the case for pictures in Experiments 1 and 2. This procedure was chosen in order to make it easier for participants to perceive subtle differences in terms of emotional contents and complexity. The order of the ratings (familiarity, complexity, pleasantness and arousal) was the same for all participants. A comparison between the inter-relationships of familiarity, arousal and pleasantness with those based on data of Experiment 1 as reported in Marin et al. [83] revealed similar results regarding the direction and strength of the relationships, thus making the presence of order effects unlikely. After the experiment, participants filled the questionnaire. All stimuli were played through an external sound card (E-MU audio

interface, 0204/USB) and Sennheiser high-quality (HD 380 pro) headphones. The volume was fixed to approximately 72 dB SPL (a-weighted), which constitutes an averaged dB value measured by a commercial dB-meter (Voltcraft SL-400 decibel meter that was calibrated immediately prior to usage) during the presentation of the second practice trial. This excerpt was one of the loudest trials in the set, and participants were asked whether they were comfortable with this loudness level before the actual experiment started. The experiment lasted around 90 minutes in total.

In order to obtain objective measures of complexity by means of data compression, the original uncompressed audio files in WAV format were transformed into one lossless compression format (Free Lossless Audio Codec, FLAC) and two common lossy compression formats (MP3 and Ogg Vorbis) by using Audacity 1.3.14 software. The 16 bit FLAC files were compressed to the maximum level 8. MP3 files were created with the following settings: bit-rate mode – pre-set extreme, 220-260 kbit/s; variable speed – standard; and channel mode – joint stereo. The quality setting for the transformation into Ogg Vorbis files was set to 8, which referred to a variable bit rate of 256 kbit/s. The file sizes of the audio files of the respective compression formats were assessed in bytes. The event density per second of each musical excerpt was analyzed by means of the MIRtoolbox [125] in Matlab.

Results and Discussion

Two female participants did not provide answers to two questions in the post-questionnaire, but their behavioral data was included in the analysis. Shapiro-Wilk normality tests showed that the data referring to participants' musical background, i.e., the role of music in participants' lives and the frequency of listening to classical music, was not normally distributed (all ps < .05). Mann-Whitney-U-tests indicated that males (n = 18) and females (n = 16) differed neither in their ratings of the role music played in their daily lives, U = 108.0, p = .195, r = -.22, nor in their frequency of listening to classical music, U = 108.5, p = .212, r = -.21. Regarding the questionnaire referring to the actual experiment, an independent t-test revealed that males (n = 18, M = 5.11, SD = .96) and females (n = 18, M = 5.06, SD = 1.16) did not significantly differ in their general liking for the piano solo music, $t(34)$ = .16, p = .877, r = .03. With respect to the difficulty of judging the complexity of the musical excerpts, there was also no significant difference between the genders, $t(34)$ = -.42, p = .678, r = .07 (males: M = 4.33, SD = 1.28; females: M = 4.50, SD = 1.10). Taken together, these results suggest that males and females had a similar musical background, especially with regard to classical music, and that both participant groups reported comparable levels of liking for the music played in the experiment and of the difficulty of judging complexity thereof. Further inspection of the musical background questionnaire indicated that only three participants had a high interest in classical music.

The main analysis was based on ratings averaged across participants for each musical excerpt. Inter-rater reliability was assessed by computing the average measure intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency [133,134]. A high degree of reliability was found for the four rating scales, as shown by the following results: familiarity (ICC(2, k) = .80, 95% CI [.74, .86]), complexity (ICC(2, k) = .96, 95% CI [.94, .97]), pleasantness (ICC(2, k) = .88, 95% CI [.84, .91]) and arousal (ICC(2, k) = .93, 95% CI [.91, .95]). ICCs computed for both genders separately revealed the following results for males: familiarity (ICC(2, k) = .62, 95% CI [.49, .72]), complexity (ICC(2, k) = .90, 95% CI [.87, .93]), pleasantness (ICC(2, k) = .73, 95% CI [.64, .80]) and arousal (ICC(2, k) = .88, 95% CI [.84, .91]). In the group of females, ICCs were generally somewhat higher than those observed for males: familiarity (ICC(2, k) = .70, 95% CI [.60, .78]), complexity (ICC(2, k) = .92, 95% CI [.89, .94]), pleasantness (ICC(2, k) = .83, 95% CI [.78, .88]) and arousal (ICC(2, k) = .86, 95% CI [.81, .90]).

After removal of several outliers prior to the analysis, a series of Shapiro-Wilk normality tests showed that most of the data of the four subjective ratings (familiarity, complexity, pleasantness and arousal) deviated significantly from a normal distribution when all participants were considered together (all ps < .122), and when males (all ps < .104) and females (all ps < .459) were considered separately. Shapiro-Wilk normality tests also revealed significant deviations from normality for the MP3 (p = .001) and the FLAC (p = .002) compression file sizes as well as for the measure of event density (p = .039) although several outliers were removed. It was thus decided to run non-parametric tests in order to investigate the relationships between these sets of variables.

Figure 3a depicts the distribution of the musical excerpts in the arousal-pleasantness space. These excerpts varied to a similar degree in arousal (min. M = 2.47, max. M = 5.22, range = 2.75) as in pleasantness (min. M = 2.83, max. M = 5.47, range = 2.64). A comparison of the current results with the findings of Experiment 1 described in [83] revealed that both distributions cover a similar part of the emotion space based on seven-point scales (pleasantness: ~ 3.0 to 5.5; arousal: 2.0 ~ 5.5). In comparison to the emotion spaces of Experiments 1 and 2 of the present study, no significant quadratic relationship between arousal and pleasantness was found for the current excerpts of piano solo music.

Table 9 shows the relationships between the four subjective measures (familiarity, complexity, pleasantness and arousal) in response to Romantic piano solo music, analyzed for females and males separately (Figure S3) and for both groups together (Figure 3). We observed a strong positive relationship for ratings of familiarity and pleasantness (r_s = .75), as was the case for the respective variables in the visual domain (see Experiments 1 and 2). Familiarity did not significantly correlate with subjective complexity judgments of music (r_s = .03), which stands in contrast to the negative relationship observed for IAPS pictures, but at the same time, is in line with the finding for representational paintings (r_s = -.07), suggesting that familiarity and complexity may be less related in artistic stimuli if effects of gender are ignored. In contrast to the findings reported in Experiments 1 and 2, familiarity did not significantly negatively correlate with arousal (r_s = -.09). This correlation was significantly different compared to the one observed with IAPS pictures, z = -2.53, p = .011.

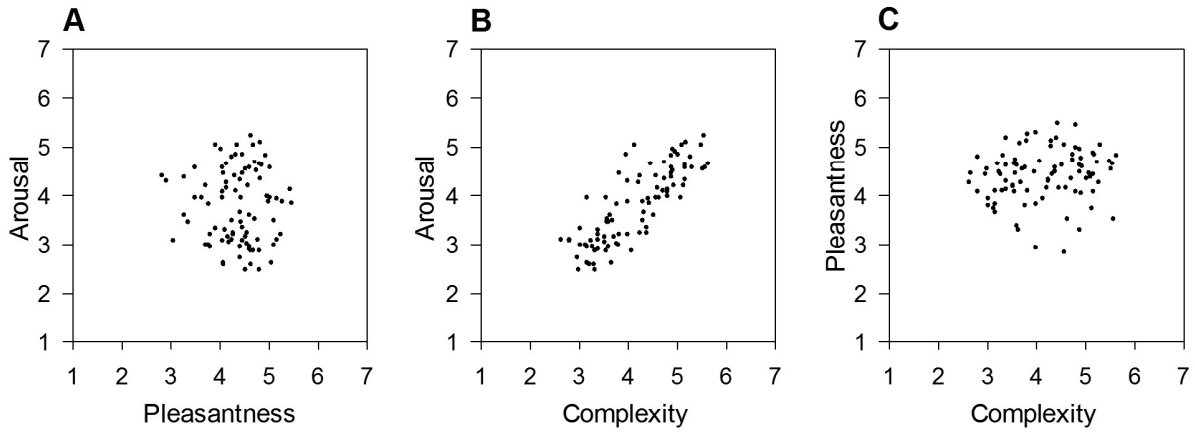

Figure 3. Relationships between pleasantness, arousal and complexity in a set of piano solo music excerpts. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Distribution of excerpts of Romantic piano solo music in the pleasantness-arousal space. B) Relationship between mean complexity and arousal ratings. C) Relationship between mean complexity and pleasantness ratings.

doi: 10.1371/journal.pone.0072412.g003

Table 9. Spearman's rank-order correlations between ratings of familiarity, complexity, pleasantness and arousal of piano solo music excerpts ($N_{All} \geq 88$, $n_{males} \geq 87$, $n_{females} \geq 85$).

Measure		Familiarity	Complexity	Pleasantness
Complexity	f	.11		
	m	-.21		
	All	.03		
Pleasantness	f	.64*	.07	
	m	.54*	.25	
	All	.75*	.18	
Arousal	f	-.13	.80*	-.15
	m	-.13	.83*	.18
	All	-.09	.83*	-.03

Note. *$p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

Partial Spearman's rank-order correlations between complexity, pleasantness and arousal were conducted to eliminate effects of familiarity (Table S7). Familiarity effects became particularly obvious in the group of males. The positive association between complexity and pleasantness gained in strength (r_s = .42), which was also the case for the relationship between arousal and pleasantness (r_s = .28). As expected, the relationship between complexity and arousal did not change when familiarity was controlled for (r_s = .82). Importantly, the underlying basis of the positive association between complexity and pleasantness in males, the only significant association observed in this series of experiments so far, is difficult to interpret. In the current experiment, the degree of complexity was not manipulated but variations in complexity were naturally present in the stimulus set under investigation. Stimuli were selected on the basis of their affective content [83]. The positive relationship between subjective complexity and arousal (controlling for effects of familiarity) was also present in Experiments 1 and 2, but in terms of the strength of this relationship the results of the three experiments clearly differ: The weakest correlation was observed in the set of IAPS pictures (r_s = .27), followed by the one in paintings (r_s = .59), and the strongest correlation was found in the current experiment for music (r_s = .83). The differences between these three correlations were significant (all *p*s < .003). All together, these results support Berlyne's hypothesis [9] of a positive association between arousal and complexity.

In comparison to Experiment 1, we found several indications of gender effects in response to representational art and music. For instance, complexity correlated positively with pleasantness in males (r_s = .42) in response to music, indicating that more complex music induced higher levels of pleasantness. This relationship was negative in females (Figure S3), leading to a non-significant association when both genders were considered together (Figure 3). While the correlation between subjective complexity and pleasantness was negative in both genders in response to IAPS pictures, a positive association as proposed by Nadal et al. [16] between these variables was only present in males in response to representational paintings and music. Furthermore, the relationship between familiarity and complexity differed in direction and magnitude with regard to representational art (negative in females, positive in males). This pattern of results was not present in the current experiment, in which a (non-significant) negative relationship was detected in males and a positive relationship in females. Consequently, the present findings support the view that gender affects particularly the relationships between familiarity, pleasantness and complexity in response to artistic stimuli, whereas the relationship between arousal and complexity appears to be less affected.

As a next step, we performed a mediation analysis in order to examine whether the indirect effect of complexity on pleasantness through arousal can also be detected in the musical domain. A bootstrapping approach was used to test the mediator model [132]. First, a mediator model (arousal) was computed. A dependent variable model (pleasantness) was computed in a second step. In both models, familiarity was specified as a covariate. All relationships between these variables were modeled as linear. The model predicting the mediator was significant and explained 69% of the variance in arousal (Table 10). Complexity had a strong significant positive impact on arousal (a = .79, p < .001). The dependent variable model was also significant and explained 55% of the variance in pleasantness. The results showed a significant direct effect of complexity on pleasantness after controlling for arousal (c' = .25, p = .005), which was not present in Experiments 1 and 2. The bootstrapped estimate of the indirect effect was negative and significant at the 95% confidence level, suggesting that pleasantness decreases linearly as complexity increases. Mediation models calculated for males and females separately revealed that this indirect effect was only significant in females.

While the results of Experiments 1 and 2 provided evidence for a complexity-arousal-pleasantness path model, which is essentially in line with the mediating role of arousal proposed by Berlyne [9], the current results also revealed a significant direct effect of complexity on pleasantness after controlling for arousal besides a significant indirect effect. In other words, the relationship between complexity and pleasantness is not completely mediated through arousal. This finding, if replicated with other musical styles, may have important implications for models of aesthetic responses that comprise different sensory domains. The current results further suggest that gender affects the mediating role of arousal in response to music, which we did not observe in the two experiments involving visual stimuli.

In line with our hypothesis, the file sizes of the three types of compression format (MP3, Ogg Vorbis and FLAC), but not those of the original WAV formats, correlated significantly positively with subjective complexity (Table 11). FLAC compression size yielded the highest correlation coefficient (r_s = .65). Compared to results of Experiments 1 and 2, the current correlations between compressed file size and subjective complexity for musical stimuli were generally stronger. Interestingly, the three types of compressed file sizes correlated even slightly (non-significantly) better with reported arousal (e.g., r_s = .73 for the FLAC compression file size) than with subjective complexity (e.g., r_s = .65 for the FLAC compression file size). The results further indicated that the size of MP3 and FLAC files correlated negatively, but less strongly, with familiarity in males (r_s ~ -.28). We also observed a significant negative correlation between the FLAC file size and pleasantness in females (r_s = -.31). These latter findings again highlight the need to consider gender in research on musical complexity in the context of empirical aesthetics.

There was also a significant positive correlation between event density and subjective complexity (r_s = .57), supporting our hypothesis. This correlation explained about 32% of the variance in subjective complexity and thus had a comparable strength to the ones observed for the correlations between the FLAC and MP3 compressed file sizes and subjective complexity. Event density also correlated positively with arousal (r_s = .42). The data further revealed a weak positive association between pleasantness and event density in males (r_s = .25). Taken together, these results are similar to those of Experiment 2, in which we also found indications that subjective ratings of paintings other than complexity weakly correlated with measures of edge detection and compressed file size. Specifically, the positive relationship between edge detection measures and arousal in the set of paintings and the one between compressed file size, event density and arousal in the set of piano solo music may be due to the stronger relationship between subjective complexity and arousal in

Table 10. Indirect effect of complexity (X) on pleasantness (Y) through arousal (M) with familiarity as a covariate, calculated for piano solo music (N_{All} = 86, n_{males} = 84, $n_{females}$ = 83).

		Model predicting arousal (M)		
		Coeff	***SE***	
Constant	f	1.63***	.39	
	m	-.37	.60	
	All	1.19*	.46	
Familiarity	f	-.25*	.11	
	m	.05	.14	
	All	-.22	.13	
Complexity (X)	f	.69***	.06	
	m	.95***	.07	
	All	.79***	.06	
Summary of model predicting M	f	R^2 = .65***		
	m	R^2 = .70***		
	All	R^2 = .69***		
		Model predicting pleasantness (Y)		
Constant	f	2.57***	.38	
	m	1.00	.48	
	All	1.23**	.39	
Familiarity	f	.73***	.11	
	m	.79***	.11	
	All	.94***	.10	
Complexity (X)	f	.23**	.08	
	m	.25*	.10	
	All	.25**	.09	
Arousal (M)	f	-.34**	.10	
	m	-.04	.09	
	All	-.21*	.09	
Summary of model predicting Y	f	R^2 = .48***		
	m	R^2 = .40***		
	All	R^2 = .55***		
		Indirect effect	**CI 95%**	
		f	-.23	-.37	-.10
		m	-.04	-.17	.11
		All	-.17	-.32	-.03

Note. X = independent variable; M = mediator variable; Y = dependent variable; *Coeff* = unstandardized coefficient; f = females; m = males; All = all participants; CI = confidence interval; p < .05, p < .01, p < .001.

Table 11. Spearman's rank-order correlations between compression algorithms applied to excerpts of piano solo music, their event density and rated familiarity, complexity, pleasantness and arousal ($N_{All} \geq 85$, $n_{males} \geq 84$, $n_{females} \geq 83$).

Measure		MP3	Ogg Vorbis	FLAC	Event Density
Familiarity	f	-.19	.01	-.18	.23
	m	-.28	-.03	-.28	.11
	All	-.20	.02	-.22	.23
Complexity	f	.55*	.35*	.58*	.54*
	m	.61*	.41*	.70*	.56*
	All	.58*	.38*	.65*	.57*
Pleasantness	f	-.26	-.13	-.31*	.15
	m	.01	.10	-.06	.25
	All	-.17	-.02	-.24	.25
Arousal	f	.61*	.38*	.69*	.37*
	m	.64*	.44*	.74*	.43*
	All	.62*	.42*	.73*	.42*

Note. * $p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

visual and musical artistic stimuli compared to IAPS pictures. Table S8 reports the inter-relationships between all objective measures of musical complexity.

A linear stepwise regression analysis was conducted to explore whether subjective or objective measures were better predictors of subjective complexity (Table 12). Familiarity was entered in the first block, followed by arousal and pleasantness as well as by the three measures of file size and event density in a second block. The final model (after four steps) was significant with an adjusted R^2 of .75, $F(4,74) = 59.21$, $p < .001$. Arousal was the strongest predictor ($\beta = .70$, $p < .001$), followed by pleasantness ($\beta = .28$, $p = .001$) and the FLAC file size ($\beta = .20$, $p = .03$). It is notable that arousal was the strongest predictor in the model, which was also the case in the model predicting subjective complexity of representational art in Experiment 2, but not in the model predicting subjective complexity ratings of IAPS pictures in Experiment 1. However, this pattern of results changed when males and females were analyzed separately. Familiarity was again entered in a first step. The model based on data from the male group was significant with an adjusted R^2 of .79, $F(5,66) = 55.41$, $p < .001$, revealing the following significant predictors: arousal ($\beta = .60$, $p < .001$), event density ($\beta = .24$, $p = .001$), MP3 file size ($\beta = .17$ $p = .022$) and pleasantness ($\beta = .14$, $p = .046$). The regression analysis based on data of the female group was similar, indicating a significant model with an adjusted R^2 of .74, $F(5,66) = 42.03$, $p < .001$, and the following predictors: arousal ($\beta = .61$, $p < .001$), event density ($\beta = .26$, $p = .001$), MP3 file size ($\beta = .19$ $p = .002$) and pleasantness ($\beta = .18$, $p = .026$). The finding that the type and order of the predictors are the same when both genders are analyzed separately and different when a model is fitted to the averaged data is worth noting. One possible explanation may relate to the different directions of the correlations between complexity and familiarity in males and females (Table 9). This result not only shows that familiarity is a crucial variable in the field of empirical aesthetics, but also that gender effects may distort findings that are based on averaged data across both groups. In other words, research designs that balance for gender without applying an in-depth analysis may yield misleading results.

Table 12. Summary of linear stepwise regression analysis for seven variables predicting subjective complexity ratings of piano solo music (N = 79).

Variable	*B*	*SE B*	B
Step 1			
Constant	4.01	.76	
Familiarity	.04	.24	.02
Adjusted R^2	-.01		
F	.03		
Step 2			
Constant	.09	.49	
Arousal	.91	.07	.85***
Familiarity	.21	.13	.10
Adjusted R^2	.71		
F	98.11***		
ΔR^2	.72		
Step 3			
Constant	-.43	.51	
Arousal	.91	.06	.85***
Pleasantness	.35	.13	.22**
Familiarity	-.12	.17	-.06
Adjusted R^2	.74		
F	73.80***		
ΔR^2	.03		
Step 4			
Constant	-1.37	.66	
Arousal	.75	.10	.70***
Pleasantness	.44	.13	.28**
FLAC file size	8.15E-6	< .001	.20*
Familiarity	-.14	.17	-.07
Adjusted R^2	.75		
F	59.21***		
ΔR^2	.02		

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; *B* = unstandardized regression coefficient; SE = standard error; β = standardized regression coefficient; ΔR^2 = difference in the proportion of variance explained.

A stepwise regression model predicting subjective complexity only by means of objective measures of complexity (Table S9) showed that FLAC file size ($\beta = .34$, $p = .001$), event density ($\beta = .46$, $p < .001$) and MP3 file size ($\beta = .27$, $p < .001$) were significant predictors, explaining approximately 60% of the variance in subjective complexity, $F(3,75) = 43.12$, $p < .001$. These results did not differ between males and females. In comparison to recent models of musical complexity based on objective measures [50,115], the current model accounts for a similar proportion of variance by using fewer predictors, which furthermore can be computed more easily.

Experiment 4

Methods

Participants. Forty German-speaking psychology students (20 males, 24.5 ± 6.0 years, age range 20-46 years; 20 females, 22.2 ± 1.7 years, age range 20-26 years) participated in the experiment in return for course credits. These participants did not take part in any other experiment in this study. All participants were enculturated in Western tonal music from their birth onwards and were non-musicians, i.e., they were neither playing a musical instrument at the time of the experiment nor had they played an instrument for longer than four years in their childhood or youth. Twenty-one participants had played an instrument in their past (musical training = 1.01 ± 1.12 years). All participants reported normal hearing.

Materials. In this experiment, the aim was to manipulate musical complexity by the number of musical instruments present in an auditory scene, that is, one instrument (piano solo music) vs. three instruments (piano trios). This approach made a meaningful comparison with Experiments 1 and 2, in which visual complexity was manipulated in a similar vein, possible. Thus, 40 piano trio excerpts (see Supporting Information Stimulus List S4) were pre-selected by the first author, a trained musicologist, from the Romantic music repertoire to fall into one of the four emotion quadrants defined by the dimensions of arousal and pleasantness [18]. The first author had experience in the selection of music excerpts in Experiment 3 which helped her to choose an appropriate range of emotions induced by piano trios. All excerpts were taken from piano trios of representative Romantic composers, such as Schubert, Schumann, Brahms, Spohr and Liszt, and in the standard piano trio instrumentation (piano, violin and violoncello). It was ensured that all three instruments were present in an excerpt and that the emotional content was not changing over time. All stimuli were cut to 25 s of length, and a 500 ms fade in/out was added to the excerpts using Audacity 1.3.14 software. The sounds were not modified in any other way. Based on the results reported in Experiment 3, 20 piano solo music excerpts were chosen as best representatives of each emotion quadrant spanned by arousal and pleasantness (see Supporting Information Stimulus L3). This guaranteed that the complete range of the emotion space of this particular musical style was covered by 40 excerpts. The same questionnaire as in Experiment 3 was used to assess participants' musical background and to collect data regarding their liking for the music and the experienced level of difficulty of judging complexity.

Procedure. In general, the experimental procedure was identical to the one described in Experiment 3, except for that participants were tested in pairs separated by a wall in a quiet room. Each participant was wearing Sennheiser (HD 380 pro) headphones. Furthermore, the piano solo music and piano trio excerpts were blocked according to their position in the arousal-pleasantness space. All trials (10 piano solo music excerpts and 10 piano Trió excerpts) were randomized within the four blocks, which were also randomized across participants. After the experiment, participants completed the questionnaire and were debriefed and dismissed. The experimental session lasted approximately 80 minutes on average. The 40 piano trio excerpts in WAV format were also transformed into MP3, Ogg Vorbis and FLAC files. The file size and event density of each musical excerpt were determined as described in Experiment 3.

Results and Discussion

A series of Shapiro-Wilk normality tests indicated that the data referring to participants' musical background, i.e., the role of music in participants' lives and the frequency of listening to classical music, was largely not normally distributed when the data was considered separately for males and females (all *p*s < .140). Similarly, the distributions of the variables regarding the general liking of the music presented in the experiment and the difficulty of giving complexity ratings were not normally distributed (all *p*s < .219). Mann–Whitney *U*-tests indicated that males (n = 20) and females (n = 20) differed neither in their ratings of the role music plays in their daily lives, U = 200.0, p = 1.00, r =.00, nor in their frequency of listening to classical music, U = 192.5, p = .836, r = -.03. Moreover, males and females liked the musical stimuli to a similar extent, U = 171.5, p = .423, r = -.13, and they also reported a similar degree of difficulty of judging the complexity of the musical excerpts, U = 188.5, p = .750, r = -.05.

Further Mann–Whitney *U*-tests revealed that participants of Experiments 3 and 4 did not differ with regard to the role music played in their lives, U = 594.0, p = .329, r = -.11, and also not with regard to the frequency of listening to classical music, U = 645.0, p = .699, r = -.04. In a similar vein, the participant groups did not significantly differ in their liking ratings for the musical stimuli, U = 654.5, p = .480, r = -.08, and in their self-reported difficulty of judging musical complexity, U = 657.0, p = .500, r = -.08. These results thus allowed a meaningful comparison between the results of the two experiments involving musical stimuli.

As in Experiments 1-3, the main analysis was based on ratings for each musical excerpt averaged across participants. Inter-rater reliability was assessed by computing the average measure intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency [133,134]. A high degree of reliability was found for the four rating scales when all participants were considered together, as shown by the following results: familiarity (ICC(2, k) = .73, 95% CI [.63, .81]), complexity (ICC(2, k) = .96, 95% CI [.95, .97]), pleasantness (ICC(2, k) = .79, 95% CI [.72, .85]) and arousal (ICC(2, k) = .93, 95% CI [.90, .95]). Similar results were obtained for the group of males, except for that inter-rater reliability was not particularly high for familiarity ratings: familiarity (ICC(2, k) = .44, 95% CI [.25, .61]), complexity (ICC(2, k) = .92, 95% CI [.89, .94]), pleasantness (ICC(2, k) = .65, 95% CI [.52, .75]) and arousal (ICC(2, k) = .87, 95% CI [.82, .91]). Inter-rater reliability was also considered sufficiently high in females: familiarity (ICC(2, k) = .62, 95% CI [.49, .73]), complexity (ICC(2, k) = .92, 95% CI [.90, .95]), pleasantness (ICC(2, k) = .69, 95% CI [.58, .78]) and arousal (ICC(2, k) = .87, 95% CI [.82, .90]). It is worth noting that although non-musicians participated in the experiment, the inter-rater reliability for complexity ratings was very high, suggesting that

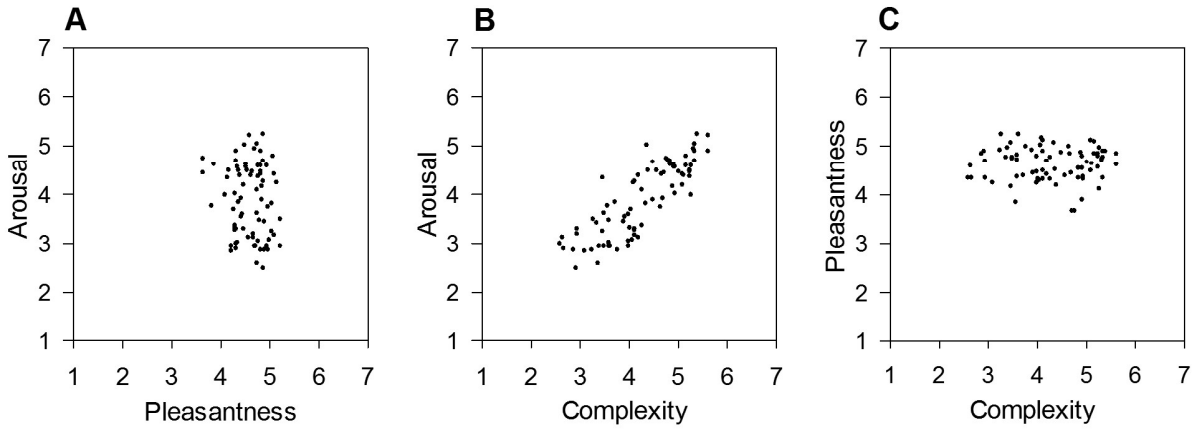

Figure 4. Relationships between pleasantness, arousal and complexity in a set of piano solo and chamber music excerpts. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Distribution of Romantic piano solo music and piano trio excerpts in the pleasantness-arousal space. B) Relationship between mean complexity and arousal ratings. C) Relationship between mean complexity and pleasantness ratings.

doi: 10.1371/journal.pone.0072412.g004

untrained listeners are able to perceive differences in musical complexity within and across musical styles of Romantic music in a reliable fashion. We also obtained high inter-rater reliability for complexity ratings of representational art in Experiment 3.

After the removal of outliers two *SD* above or below the mean, Shapiro-Wilk normality tests showed that most of the subjective ratings with regard to the musical excerpts, with the exception of those of familiarity, were not normally distributed when both genders were considered together (all *p*s < .03), and when males (all *p*s < .10) and females (all *p*s < .15) were analyzed separately. Deviations from normality were also observed for the Ogg Vorbis compressed file size ($p < .001$), therefore it was decided to conduct non-parametric correlation analyses as in Experiments 1-3. A comparison between the distributions of musical excerpts in the arousal-pleasantness space in Experiment 3 (Figure 3a) and 4 (Figure 4a) revealed that excerpts were less spread apart in terms of pleasantness when piano trios and piano solo music were combined in one stimulus set.

Here, complexity was manipulated by mixing piano solo excerpts with piano trios on the grounds that these have a greater number of melodic parts than the piano solo excerpts used in Experiment 3. However, adding additional melodic parts does more than simply add extra complexity to a piece of music. It may also add timbral streaming, harmonic richness and complexity, and greater expressive scope. It thus remains possible that these factors were responsible for any differences found between Experiments 3 and 4, rather than differences in subjective complexity per se. A comparison between the subjective complexity ratings of the original 40 piano solo and 40 piano trio excerpts revealed that piano trio excerpts (M = 4.57, SD = .68) were indeed rated as more complex than piano solo excerpts (M = 3.93, SD = .65), $t(78) = -4.34$, $p < .001$, r = .44.

The inter-relationships between the four subjective variables (Table 13 and Figure 4) were generally similar to the ones observed for piano solo music in Experiment 3 (Table 9) when familiarity was not considered in the analysis, thus suggesting that the manipulation of complexity by changing the audible number of instruments in an auditory scene had a small effect on these inter-relationships. For instance, natural variations in complexity in piano solo music were strongly correlated to degrees of felt arousal, which was also the case when complexity was deliberately varied (r_s = .84). The findings further indicated significant positive correlations between familiarity and pleasantness in all three groups. Furthermore, the pattern of correlations between arousal and pleasantness was similar in direction as in Experiment 3. Specifically, the current results showed that arousal and pleasantness were negatively associated in females (r_s = -.29) but positively in males (r_s = .22), yielding a slightly negative correlation on average. Another notable finding concerns the positive relationship between complexity and pleasantness in males (r_s = .22) and a concurrent negative relationship in females (Figure S4), whereas in Experiment 3 complexity and pleasantness were not related in females. Moreover, the negative correlation between familiarity and arousal was stronger when all participants were considered in the analysis (r_s = -.24) compared to results presented in Experiment 3. In a similar vein, there were clear indications for a negative correlation between familiarity and complexity in males and females, which was not the case for females in Experiment 3.

Partial Spearman's rank-order correlations controlling for effects of familiarity were computed in a next step (Table S10), and the results were more or less identical to the findings of Experiment 3, indicating that the observed differences in the inter-relationships between the subjective variables described above were not due to the different stimulus sets. Males showed a significant positive correlation between complexity and pleasantness (r_s = .41), whereas this relationship remained negative but not significant in females (r_s = -.14). The relationship between arousal and pleasantness got stronger

Table 13. Spearman's rank-order correlations between ratings of familiarity, complexity, pleasantness and arousal in response to set of piano solo and piano trio excerpts ($N_{All} \geq 76$, $n_{males} \geq 74$, $n_{females} \geq 75$).

Measure		Familiarity	Complexity	Pleasantness
Complexity	f	-.12		
	m	-.18		
	All	-.21		
Pleasantness	f	.61*	-.26	
	m	.42*	.22	
	All	.67*	.02	
Arousal	f	-.15	.83*	-.29
	m	-.10	.76*	.22
	All	-.24	.84*	-.10

Note. * $p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

only for males (r_s = .35) but not for females (r_s = -.23) when familiarity was considered in the analysis. The association between complexity and arousal also remained unchanged in all analyses after controlling for familiarity effects. In summary, it can be concluded that familiarity plays an indispensable role in the investigation of the inter-relationships between complexity, arousal and pleasantness. Most importantly, it seems that the relationship between complexity and pleasantness cannot be meaningfully discussed without considering gender. Future studies may investigate the underlying reasons of these gender effects and whether they hold true for other musical styles.

As in Experiments 1-3, a mediation analysis based on a bootstrapping approach [132] was performed in order to examine the presence of an indirect effect of complexity on pleasantness through arousal. Such an indirect effect would be expected according to Berlyne's collative-motivation model [9]. First, a mediator model (arousal) was computed, and then a dependent variable model (pleasantness) was computed in a second step. Familiarity was specified as a covariate in both models. All relationships between the variables were modeled as linear. In general, the current results were very similar to those of Experiment 3. The model predicting the mediator was significant and explained 70% of the variance in arousal (Table 14). As expected, complexity had a strong significant positive impact on arousal (a = .76, p < .001). The dependent variable model was also significant and explained 54% of the variance in pleasantness. Furthermore, the results showed a significant direct effect of complexity on pleasantness after controlling for arousal (c' = .20, p = .003), which we did not find in Experiments 1 and 2. The bootstrapped estimate of the indirect effect was negative and significant at the 95% confidence level, suggesting that pleasantness decreases linearly as complexity increases through its effect on arousal. As in Experiment 3, mediation models calculated for males and females separately revealed that this indirect effect was only significant in females, although both groups had similar musical listening backgrounds. We consider the replication of the results of

Table 14. Indirect effect of complexity (X) on pleasantness (Y) through arousal (M) with familiarity as a covariate, calculated for excerpts of piano solo music and piano trios (N_{All} = 76, n_{males} = 73, $n_{females}$ = 72).

		Model predicting arousal (M)			
		Coeff	*SE*		
Constant	f	.49	.57		
	m	.007	.78		
	All	.83	.61		
Familiarity	f	.003	.15		
	m	.26	.22		
	All	-.07	.16		
Complexity (X)	f	.74***	.06		
	m	.76***	.07		
	All	.76***	.06		
Summary of model predicting M	f	R^2 = .69***			
	m	R^2 = .63***			
	All	R^2 = .70***			
		Model predicting pleasantness (Y)			
Constant	f	2.74***	.39		
	m	1.00	.57		
	All	1.83***	.37		
Familiarity	f	.73***	.10		
	m	.91***	.16		
	All	.84***	.10		
Complexity (X)	f	.15*	.07		
	m	.22*	.08		
	All	.20**	.06		
Arousal (M)	f	-.24**	.08		
	m	-.05	.09		
	All	-.16*	.07		
Summary of model predicting Y	f	R^2 = .48***			
	m	R^2 = .36***			
	All	R^2 = .54***			
		Indirect effect	**CI 95%**		
		f	-.18	-.31	-.05
		m	-.04	-.21	.08
		All	-.12	-.24	-.01

Note. X = independent variable; M = mediator variable; Y = dependent variable; *Coeff* = unstandardized coefficient; f = females; m = males; All = all participants; CI = confidence interval; p < .05, p < .01, p < .001.

Experiment 3, with partially different musical stimuli and a new sample of participants, to be a noteworthy finding. First, the observed gender effects with regard to the mediation effect seem to be robust when different styles of Romantic music are used as stimulus material. Second, the co-occurrence of significant direct and indirect effects may be a particularity of the musical domain because direct effects were not present in the current two experiments involving visual stimuli. Future studies employing different types of visual and musical stimuli under a common experimental framework may thus reveal whether the current findings can be generalized.

The Spearman's rank-order correlations between subjective and objective complexity as measured by compressed file size (Table 15) revealed a positive association between all four

Table 15. Spearman's rank-order correlations between compression algorithms applied to excerpts of piano solo and piano trio music, their event density and rated familiarity, complexity, pleasantness and arousal ($N_{All} \geq 71$, $n_{males} \geq 70$, $n_{females} \geq 71$).

Measure		MP3	Ogg Vorbis	FLAC	Event Density
Familiarity	f	-.15	.02	-.12	.24
	m	-.18	-.10	-.23	.06
	All	-.22	-.13	-.22	.21
Complexity	f	.66*	.62*	.79*	.37*
	m	.63*	.55*	.70*	.47*
	All	.66*	.60*	.78*	.43*
Pleasantness	f	-.27	-.11	-.20	.20
	m	.09	.25	.03	.34*
	All	-.06	.06	-.07	.29
Arousal	f	.64*	.58*	.80*	.15
	m	.65*	.60*	.70*	.26
	All	.64*	.59*	.78*	.20

Note. * $p < .05$ after Bonferroni-Holm correction; f = females; m = males; All = all participants; the *dfs* are not the same for all correlations due to slightly different numbers of outliers.

variables. In comparison to results of Experiment 3, the magnitude of these correlations was somewhat larger, with the exception of event density (r_s = .43), when piano trios and piano solo music were combined in one stimulus set. As in Experiment 3, the highest correlation was observed for the compression size based on the FLAC format (r_s = .78), explaining around 60% of the variance in the data. Bearing in mind that Streich's model of musical complexity [50] reached a very similar performance with twelve complexity descriptors (around 60% without considering variables based on compression file size), this result clearly highlights the usefulness of compression file size in future models of musical complexity. The finding that event density was a stronger predictor of subjective complexity in Experiment 3 than in the present stimulus set may relate to the measure itself and how it is calculated. It is likely that the current event density algorithm performs worse when several instruments are present in an auditory scene. Consequently, this may point to a possible advantage of measures based on compressed file size over measures based on audio feature extraction, especially in cases in which several musical styles are mixed together. However, several measures based on feature extraction will need to be considered in future studies in order to make such a claim valid.

Moreover, as observed in Experiment 3, the compressed file size measures and event density did not correlate solely with subjective complexity. The strength of the relationships between compressed file size and arousal were similar for the two data sets, whereas event density did not significantly correlate with arousal in Experiment 4. Interestingly, indications for a negative correlation between familiarity and the MP3 and FLAC file sizes were present in Experiments 3 and 4. In contrast to these findings, familiarity and event density correlated positively with each other in both experiments. In addition, the current data revealed negative correlations between pleasantness and the three types of compression file size in females, which was also present in Experiment 3. In Experiment 4, we further observed a significant positive correlation between event density and pleasantness in males (r_s = .34), an association which was already weakly present in all groups in Experiment 3. A summary of the inter-relationships between the objective measures of musical complexity applied to piano solo and piano trio excerpts is shown in Table S8.

Linear stepwise regression analyses were calculated to examine how subjective ratings of familiarity, arousal and pleasantness as well as objective measures of complexity explain variance in subjective complexity. In the model based on all participants (Table 16), familiarity was entered as a first step followed by six predictors in a second step. In contrast to findings of Experiment 3, familiarity was a significant predictor of subjective complexity (β = -.25, p = .048). The final model after four steps (adjusted R^2 = .80, $F(4,59)$ = 62.52, p < .001) revealed three other significant predictors, namely arousal (β = .68, p < .001), event density (β = .24, p < .001) and MP3 file size (β = .17, p = .024). Separate regression analyses computed for both genders showed very similar results. The model for males, adjusted R^2 = .74, $F(4,56)$ = 42.61, p < .001, yielded the following four predictors: familiarity (β = -.11, p = .126), arousal (β = .57, p < .001), event density (β = .21, p = .004) and MP3 file size (β = .26, p = .006). The pattern of results only changed minimally for females because MP3 file size did not significantly explain variance in subjective complexity. The model, adjusted R^2 = .77, $F(3,59)$ = 69.58, p < .001, was based on the following predictors: familiarity (β = -.12, p = .066), arousal (β = .77, p < .001), and event density (β = .28, p < .001). Importantly, a similar pattern of results was observed in Experiment 3 when the groups of males and females were analyzed separately, with the exception that pleasantness did not emerge as a significant predictor of subjective complexity in the current models. Finally, a model of subjective complexity solely based on objective complexity measures was built (Table S11). The final model, adjusted R^2 = .67, $F(3,63)$ = 44.77, p < .001, comprised three predictors: FLAC file size (β = .55, p < .001), MP3 file size (β = .28, p = .003) and event density (β = .21, p = .008). A comparison with results of Experiment 3 (Table S9) suggests that the significant predictors were the same for both types of stimulus sets. However, event density was a weaker predictor of subjective complexity when piano solo and piano trio excerpts were mixed in one stimulus set. The pattern of results did not change when models were computed separately for males and females.

General Discussion

Research on subjective complexity and its relationship to hedonic value has been of substantial interest to the empirical aesthetics research community and beyond. However, the emotional content of the stimuli under investigation has been largely ignored [74], which may have limited the interpretations and generalizability of results. Here, we presented a cross-domain approach to the study of the relationship between

Table 16. Summary of linear stepwise regression analysis for seven variables predicting subjective complexity ratings of piano solo and piano trio excerpts (N = 64).

Variable	*B*	*SE B*	β
Step 1			
Constant	6.32	1.03	
Familiarity	-.68	.34	-.25*
Adjusted R^2	.05		
F	4.10*		
Step 2			
Constant	.77	.71	
Arousal	.99	.08	.85***
Familiarity	-.10	.19	-.04
Adjusted R^2	.72		
F	84.79***		
ΔR^2	.67		
Step 3			
Constant	.97	.64	
Arousal	.90	.07	.77***
Event density	.20	.05	.25***
Familiarity	-.27	.17	-.10
Adjusted R^2	.78		
F	76.04***		
ΔR^2	.06		
Step 4			
Constant	-2.26	1.53	
Arousal	.79	.09	.68***
Event density	.19	.05	.24***
MP3 file size	4.72E-6	< .001	.17*
Familiarity	-.23	.17	-.08
Adjusted R^2	.80		
F	62.52***		
ΔR^2	.02		

Note. * $p < .05$, ** $p < .01$, ***$p < .001$, *B* = unstandardized regression coefficient; SE = standard error; β = standardized regression coefficient; ΔR^2 = difference in the proportion of variance explained.

subjective complexity, felt pleasantness and arousal in affective stimuli sets of environmental scenes (IAPS pictures), representational paintings, and music. Russell's circumplex model of affect [18] served as a common framework for the selection of representative stimuli inducing a wide range of emotions. Furthermore, we also applied recent advances in computational measures of complexity to the prediction of subjective complexity in the visual domain, and most importantly, introduced these methodological approaches to the musical domain for the first time. In this endeavor, complexity, regarded as a multidimensional concept, was varied by changing the number of elements present in a visual or musical scene. Last, the possible role of gender in these relationships was explored.

In all four experiments, the current data provided convincing evidence for a positive association between subjective complexity and arousal, which is in line with Berlyne's collative-motivation model [7,10]. When controlling for effects of familiarity, the weakest association was found for environmental scenes (r_s = .27), followed by representational paintings (r_s = .59) and music (r_s = .84). In the present study, environmental scenes and representational paintings were matched for semantic contents and for average brightness; therefore, it can be argued that the differences in strength of this relationship may not only refer to the type of modality but also to the degree of artistic quality and motivational relevance. In other words, for artworks we observed a stronger relationship between complexity and arousal. Future studies may examine this hypothesis by comparing different types of visual and auditory stimuli varying in their artistic quality.

Based on recent findings by Nadal et al. in the visual domain [16], that essentially argue for a multidimensional nature of complexity, we predicted a positive association between subjective complexity and pleasantness because complexity was manipulated by the number of elements present in a visual or auditory scene. This hypothesis stands in stark contrast to Berlyne's proposed inverted-U relationship between these variables [7,9]. In general, we found indications of a negative association between complexity and pleasantness for both genders when viewing environmental scenes, but this relationship was modulated by gender when viewing representational paintings or listening to music. Furthermore, we did not find support for a positive association between subjective complexity and pleasantness for the four types of visual and musical stimuli when effects of familiarity were controlled for in the analysis. Nevertheless, analyzing the data for both genders separately revealed that subjective complexity and pleasantness are indeed positively associated in males in both experiments involving musical stimuli (piano solo music: r_s = .42; piano solo music combined with piano trios: r_s = .41) but not in females. It needs to be seen whether this result holds for different musical styles. These findings may suggest that there is no common trend for the relationship between complexity and pleasantness in the current framework of the present cross-domain study, although this was clearly the case for the relationship between complexity and arousal. Further experiments may explore whether these findings are due to the specific emotion model on which the selection of stimuli was based [18], or whether they can be replicated within the context of other affective approaches or with other measures of hedonic value, such as preference and beauty. Specifically, based on our results it can be surmised that not only complexity, but also hedonic value, should be treated as a multidimensional concept. However, this hypothesis would require a more thorough investigation of the complexity-hedonic value relationship, considering effects of familiarity and gender alike.

Our results also shed light on the question of whether the complexity-pleasantness relationship is mediated by arousal as proposed by Berlyne [7,10]. By applying mediation analyses based on a bootstrapping approach [132] we observed indirect effects of complexity on pleasantness through arousal in all four experiments, which lends support to Berlyne's theory [7,10]. In the visual domain, arousal was identified as a mediator in response to affective environmental scenes. No direct effect of complexity on pleasantness was identified in this case. The findings for representational paintings somewhat

differed because the indirect effect was only identified when males and females were analyzed separately, probably due to differential effects of familiarity. The indirect effect was only significant for paintings moderate and high in complexity but not for paintings of low complexity. As was the case in the experiment involving environmental scenes, we did not find indications of a direct effect of complexity on pleasantness with regard to representational paintings. In the musical domain, the path model investigating the complexity-arousal-pleasantness relationship revealed that arousal was a significant mediator only for female participants. Moreover, a significant direct effect of complexity on pleasantness was present, which was not the case in the visual domain. These findings may point towards a potentially interesting cross-domain difference between the underlying cognitive and neural mechanisms of the relationship between complexity, arousal and pleasantness when aspects of familiarity are controlled for. However, further studies are necessary to replicate these findings with different types of visual and musical stimuli, as well as with different types of hedonic measures.

If one considers adopting an affective approach to the study of complexity, including the concept of arousal, there are several issues that warrant attention: Arousal presumably may not form a uniform concept as described in Berlyne's psychobiological model; instead, current behavioral and neuroscientific findings suggest that there may be different forms of affective arousal with related multiple underlying brain mechanisms. For example, energy arousal has been differentiated from tension arousal [142]. In line with this argument, a dissociation between the direction of change of different types of behavioral, psychophysiological and neural measures of arousal in response to a stimulus has been reported [58,143,146], which may be partly explicable by ongoing emotion regulation processes [147]. In spite of these potential difficulties associated with the incorporation of arousal in the study of complexity, the affect-as-information hypothesis (for a review, see 147) and the attributed role of arousal in evaluative judgments, cognitive processing and memory [148] strongly support the view that aesthetic theories ignoring the role of arousal may be incomplete [51,52,149]. Indeed, arousal has been identified as a predictor of preference in the visual [150] and musical [151,152] domains outside the context of Berlyne's theory [9].

Measures based on compressed file size and edge detection algorithms have recently been applied to the study of visual complexity [74,102,104]. One of the main findings of the present study is that when complexity is manipulated through the number of elements present in a visual scene, certain measures of objective complexity are more suitable for environmental scenes than for visual art and vice versa. Importantly, this is not in line with Forsythe et al. [74], who found no significant differences for non-affective environmental scenes and representational paintings when exploring the relationships between subjective complexity and objective measures, such as JPEG and GIF compressed file size and perimeter detection, respectively. However, Forsythe et al. [74] did not vary complexity along one specific dimension in their stimuli set, which limits the comparison with our current results. Nevertheless, our findings indicate that the performance of compression file size and different measures of edge detection may depend on how complexity is manipulated in the visual materials.

Specifically, the findings obtained in the present study showed that compressed file size only correlated positively with subjective complexity judgments of photographs of environmental scenes but not with complexity judgments of representational paintings. TIFF (r_s = .53) and JPEG (r_s = .52) compression file sizes yielded the strongest correlations with subjective complexity ratings of environmental scenes. In contrast to Forsythe et al. [74], the GIF compression format yielded the lowest correlation (r_s = .29). The moderate correlation between JPEG compression size and subjective complexity was also reported by Cavalcante et al. [104], suggesting that JPEG may be a reliable measure of objective complexity across different types of stimulus sets of environmental scenes and taken as a proxy for subjective ratings. However, none of the four compression file formats correlated significantly with subjective complexity judgments of representational paintings. At first sight, this finding is rather surprising since it is not in line with previous reports [74]. Nevertheless, it may be explained by the way complexity was manipulated in the current stimuli sets. A uniform background perceived as simple may be easily compressible in photographs of environmental scenes but not in representational paintings, in which the structure of the background does not necessarily contain a high amount of redundant information. For example, some backgrounds in paintings of simple figure-ground compositions may consist of a large number of individual brush strokes that form a rather uniform background from a subjective perspective. Compared to a background of a uniform blue sky in a photograph, the background in a painting may be less compressible although it perceptually appears similarly simple. In brief, these findings suggest that digitized versions of visual art cannot be treated like photographs of environmental scenes in the discussion about the relationship between subjective complexity and compressed file size when other factors, such as semantic and affective contents, brightness, picture size, presentation time and the kind of complexity manipulation are controlled for.

In a similar vein, disparities between environmental scenes and representational paintings were noted when comparing the performance of a set of edge detection measures to predict subjective complexity. For environmental scenes, the mean and alpha RMS contrast measures yielded the strongest correlations (r_s ~ .59), followed by perimeter detection (r_s ~ .45) and the standard deviation of the RMS contrast measure (r_s = .43). The mean and beta phase congruency measures worked reasonably well with a correlation of r_s ~ .35. The pattern of results looks somewhat different for representational paintings. The standard deviation of the phase congruency measure performed best (r_s = .38), followed by the beta measure of phase congruency (r_s = .30), and the alpha and standard deviation of the RMS contrast measures (r_s ~ .25). Taken together, these results suggest that the correlation strength between edge detection measures and subjective complexity is stronger for photographs of environmental scenes than for

paintings. This finding is in line with Forsythe et al. [74], who reported a stronger positive correlation for environmental scenes (r_s = .54) than for representational paintings (r_s = .37) with respect to perimeter detection. In the present study, perimeter detection did not yield correlations as strong as those reported by Forsythe et al. [74] since this measure only performed moderately well for environmental scenes but not for paintings.

Furthermore, the current results replicate and extend those by Cavalcante et al. [104], who found that RMS contrast measures outperformed the JPEG compression format in predicting subjective complexity. Our data also revealed a positive correlation of r_s ~ .60 for environmental scenes and a weaker association in the set of representational paintings. This finding indicates that RMS contrast measures perform better or equally well as measures such as perimeter detection and Canny edge detection, depending on the stimuli type under investigation. In a similar vein, edge detection based on phase congruency [123,153] provided satisfying correlations with subjective complexity for both environmental scenes and paintings. This measure has not been applied to the prediction of subjective complexity earlier, but our results clearly demonstrate its usefulness. The standard deviation of the phase congruency measure was the best correlate for subjective complexity judgments of paintings (r_s = .38), and the beta phase congruency measure worked similarly well for environmental scenes (r_s = .36). Further research may elucidate why these new measures of RMS contrast and phase congruency work better to predict subjective complexity for representational stimuli than previously used measures of edge detection.

It is interesting to note that we observed a negative correlation between the Canny edge detection measure and subjective complexity for representational paintings, (r_s = -.23). Canny edge detection is particularly sensitive to fine lines in an image; therefore, lines associated with brush strokes in a background of a simple figure-ground composition are detected, yielding a negative correlation with subjective complexity judgments. For environmental scenes, this correlation was positive and together with the correlation for the entropy measure it was among the weakest when the whole set of edge detection measures was considered. This stands in contrast to earlier studies which reported moderate positive correlations between subjective complexity and the Canny algorithm for different types of stimulus sets other than environmental scenes [102].

Previously it has been argued that objective measures of complexity are not related to subjective familiarity judgments of visual materials [102]. Our results generally support this finding. Nevertheless, we found indications of weak correlations between familiarity and measures of edge detection in environmental scenes as well as between familiarity and compression file sizes in representational paintings. Moreover, subjective arousal correlated with several edge detection measures (RMS contrast, phase congruency and Canny algorithm) in representational paintings. The latter finding may be due to a significant correlation between subjective complexity and arousal, which was more pronounced in the set of representational paintings compared to environmental scenes. With this in mind, it can be concluded that the use of edge detection measures can also be extended to predict other subjective measures besides subjective complexity and pleasantness in the future [74]. Conversely, compressed file size as a measure of complexity was not significantly associated with subjective ratings of arousal and pleasantness. This finding may be understood in light of the two types of objective complexity measures examined in this study: Compressed file size represents a rather abstract measure of information contents, whereas edge detection measures are more concrete and object-related.

The current study convincingly demonstrated that advances in research on the relationship between subjective and objective complexity in the visual domain can be successfully applied to the musical domain. In two experiments, the ratio between the original and the compressed file size was shown to be an appropriate correlate of subjective complexity. One experiment investigated this relationship between subjective and objective complexity in a set of Romantic piano solo music excerpts, whereas another experiment used a set of musical excerpts in which musical complexity was varied by the number of instruments present in an auditory scene, i.e., by adding excerpts of piano trio music to those of piano solo music. In fact, the obtained positive correlations were generally stronger than the ones observed in the visual domain, and this regardless of whether complexity was manipulated or naturally present in a set of affective musical stimuli.

In analogy to edge detection measures, we also computed the event density of musical excerpts as a potential correlate of subjective complexity judgments and obtained moderate correlations for the set of piano solo music and the set of piano trios combined with piano solo music. However, event density yielded stronger correlations when only one musical genre was considered. It is important to note that compressed file size and event density correlated better with subjective complexity than any of the 12 predictors of Streich's model of musical complexity [50], in which the best predictor yielded a moderate correlation of r = .42. Further research may thus investigate the potential application of compressed file size as a predictor of subjective complexity, as well as the one of event density, in stimuli sets comprising several musical styles.

Moreover, compressed file size of musical excerpts correlated strongly with subjective arousal, a finding which was also observed with regard to edge detection measures of representational paintings. It can be conjectured that the relationship between arousal and measures of objective complexity may be specific to artistic stimuli. Compressed file size and event density further correlated with familiarity and pleasantness in music, but these correlations were much weaker than the ones observed with arousal and sometimes only present in one gender. Nevertheless, these results clearly demonstrate that measures of objective complexity can be used to predict other aesthetic or affective responses in the musical domain, which extends the application of these measures. For instance, complexity measures, such as compressed file size and event density, could be integrated

into current attempts of modeling emotional responses to music [93,154].

Linear stepwise regression models controlling for effects of familiarity revealed that objective measures of complexity were better predictors of subjective complexity than subjective arousal only in the case of environmental scenes. In the visual domain, the models were moderately successful and yielded adjusted R^2 s between .40 and .51 with two or three predictors, whereas regression models of musical complexity were able to predict around 80% of the variance in subjective complexity with models based on four predictors. In general, predictive modeling employing only objective complexity measures without controlling for familiarity effects was much less successful. With regard to environmental scenes and representational paintings an adjusted R^2 of .25 was reported, while in the musical domain higher values of .62 and .67 were reached. In other words, these results suggest that the current approach taken in predicting subjective complexity in the musical domain is very promising, especially since earlier models of musical complexity proposed by Streich [50] and Mauch and Levy [115] achieved similar results by following a computationally more demanding method.

The multimodal nature of perception, and especially interactions between the visual and auditory domains, has been recently highlighted in the study of sensory processing [155,156]. While affective interactions between the musical and visual domain have already been investigated [83,157,158], direct comparisons between the processing of visual and musical complexity are scarce. Our findings that conceptually similar objective measures of complexity (compression file size and event/object density) can be successfully applied to both the visual and musical domain are in line with results reported by Boon et al. [112]. Boon et al. [112] conducted computational analyses of fractal dimension and entropy of abstract art works by Jackson Pollock and music by Johann Sebastian Bach, showing that the fractal nature of art is present across domains. Taken together, these findings add to the discussion whether the processing of complexity in the visual and musical domains, although at first sight so different in nature, should be studied within a common framework in the future.

Another noteworthy finding of the present study was the observation of gender effects. We found substantial evidence for gender effects regarding the inter-relationships of the subjective ratings of familiarity, complexity, pleasantness and arousal and for the relationships of these variables with measures of objective complexity in both the visual and musical domain. Interestingly, gender effects were more frequent in the three experiments using artistic stimuli compared to the experiment involving environmental scenes. These effects not only concerned differences with respect to the magnitude of an association, but, more importantly, also the direction of a relationship, i.e., positive versus negative. In general, this is in line with previous research on emotional and aesthetic responses that has shown differential effects of gender in the visual [19,83,85,86,139,140], and musical [20] domains. The reasons for gender effects in affective and aesthetic responses may be manifold and related to biological and sociocultural factors. For example, gender effects may be due to females' hypersensitivity to negative affective stimuli [19,20,87,159], or be explicable by the empathizer-systemizer theory [160] and the hunter-gatherer hypothesis [161]. Equally important, gender differences have been reported in the study of low-level perceptual processes, such as those involved in color perception [162], which in turn may influence the perception of affective visual stimuli. Further cross-domain research is clearly warranted to elucidate the role of gender in affective and aesthetic responses.

Conclusions

Although emotions play a crucial role in the processing of visual and musical stimuli in everyday life, subjective complexity and its relation to hedonic value have not been extensively investigated within the context of affective stimuli and emotion models so far. By comparing responses to affective environmental scenes, representational paintings and music, we found cross-domain evidence for a positive relationship between subjective complexity and arousal, which is in line with Berlyne's collative-motivation model [7,9]. In general, the data did not reveal a significant association between subjective complexity and pleasantness for affective stimuli when effects of familiarity were controlled for, which supports neither the predictions made by Berlyne [7,9] nor those of Nadal et al. [16]. A significant positive association between complexity and pleasantness was only observed in males with regard to two types of musical stimuli used in this study.

Recent advances in the study of objective complexity in the visual domain have been extended by incorporating new computational measures of compression file size and edge detection. Results indicated that the performance of these measures in predicting subjective complexity depends on the stimulus type (environmental scenes vs. representational paintings) when complexity is manipulated by the number of elements present in a visual scene. We also demonstrated that compressed file size and event density are useful predictors of subjective musical complexity in two sets of musical stimuli. Furthermore, we found instances of weak correlations between measures of objective complexity with familiarity pleasantness and arousal ratings, and these correlations seem partly to depend on gender and stimulus type. In summary, these findings are consistent with the view that only a cross-domain approach may lead to the development of a general theory about the relationship between emotion and complexity that incorporates aesthetic quality in its framework.

Acknowledgements

The authors would like to thank Bruno Gingras and Marcos Nadal Roberts for valuable comments on earlier versions of the manuscript.

Supporting Information

Figure S1. Relationships between pleasantness, arousal and complexity, analyzed for males and females, in a set of

IAPS pictures. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Relationship between pleasantness and arousal for females. B) Relationship between complexity and arousal for females. C) Relationship between complexity and pleasantness for females. D) Relationship between pleasantness and arousal for males. E) Relationship between complexity and arousal for males. F) Relationship between complexity and pleasantness for males.
(TIF)

Figure S2. Relationships between pleasantness, arousal and complexity, analyzed for males and females, in a set of representational paintings. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Relationship between pleasantness and arousal for females. B) Relationship between complexity and arousal for females. C) Relationship between complexity and pleasantness for females. D) Relationship between pleasantness and arousal for males. E) Relationship between complexity and arousal for males. F) Relationship between complexity and pleasantness for males.
(TIF)

Figure S3. Relationships between pleasantness, arousal and complexity, analyzed for males and females, in a set of piano solo music excerpts. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Relationship between pleasantness and arousal for females. B) Relationship between complexity and arousal for females. C) Relationship between complexity and pleasantness for females. D) Relationship between pleasantness and arousal for males. E) Relationship between complexity and arousal for males. F) Relationship between complexity and pleasantness for males.
(TIF)

Figure S4. Relationships between pleasantness, arousal and complexity, analyzed for males and females, in a set of piano solo and piano trio excerpts. Low numbers refer to low ratings of pleasantness, arousal and complexity, respectively. A) Relationship between pleasantness and arousal for females. B) Relationship between complexity and arousal for females. C) Relationship between complexity and pleasantness for females. D) Relationship between pleasantness and arousal for males. E) Relationship between complexity and arousal for males. F) Relationship between complexity and pleasantness for males.
(TIF)

Table S1. Partial Spearman's rank-order correlations between ratings of complexity, pleasantness and arousal, controlling for effects of familiarity, in response to IAPS pictures (*N* = 96). Delete this.
(PDF)

Table S2. Spearman's rank-order correlations between a representative set of measures of objective complexity applied to a set of IAPS pictures (*N* = 96).
(PDF)

Table S3. Summary of linear stepwise regression analysis for seven variables of objective complexity predicting subjective complexity ratings of IAPS pictures (*N* = 72).
(PDF)

Table S4. Partial Spearman's rank-order correlations, controlling for effects of familiarity, between ratings of complexity, pleasantness and arousal in response to representational paintings (*N* = 96).
(PDF)

Table S5. Spearman's rank-order correlations between a representative set of measures of objective complexity applied to a set of representational paintings (*N* =96).
(PDF)

Table S6. Summary of linear stepwise regression analysis for seven variables of objective complexity predicting subjective complexity ratings of representational paintings (*N* = 64).
(PDF)

Table S7. Partial Spearman's rank-order correlations controlling for effects of familiarity, between ratings of complexity, pleasantness and arousal in response to piano solo music (*N* = 92).
(PDF)

Table S8. Spearman's rank-order correlations between four measures of objective complexity applied to a set of piano solo excerpts (*N* = 92) and to a set of piano solo music excerpts combined with piano trio excerpts (*N* = 80).
(PDF)

Table S9. Summary of linear stepwise regression analysis for four variables of objective complexity predicting subjective complexity of piano solo music (*N* = 79).
(PDF)

Table S10. Partial Spearman's rank-order correlations, controlling for effects of familiarity, between ratings of complexity, pleasantness and arousal in response to piano solo and piano trio excerpts (*N* = 80).
(PDF)

Table S11. Summary of linear stepwise regression analysis for four variables of objective complexity predicting subjective complexity ratings of piano solo and piano trio excerpts (*N* = 67).
(PDF)

Stimulus List S1. List of 96 IAPS pictures.

(PDF)

Stimulus List S2. List of 96 representational paintings.
(PDF)

Stimulus List S3. List of 92 Romantic piano solo music excerpts.
(PDF)

Stimulus List S4. List of 40 Romantic piano chamber music excerpts.
(PDF)

Author Contributions

Conceived and designed the experiments: MMM HL. Performed the experiments: MMM. Analyzed the data: MMM. Contributed reagents/materials/analysis tools: MMM HL. Wrote the manuscript: MMM HL.

References

1. Kaplan S, Kaplan R (1982) Cognition and environment: Functioning in an uncertain world. New York: Praeger.
2. Schooler C (1984) Psychological effects of complex environments during the life span. A Rev Theory Intelligence 8: 259-281.
3. Bradley MM, Houbova P, Miccoli L, Costa VC, Lang PJ (2011) Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. Psychophysiology 48: 1543-1522. PubMed: 21649664.
4. Donderi DC (2006) Visual complexity. A Rev Psychological Bulletins 132: 73-97. doi:10.1037/0033-2909.132.1.73.
5. Gibson E (1998) Linguistic complexity. Locality of syntactic dependencies. Cognition 68: 1-76. doi:10.1016/S0010-0277(98)00035-3. PubMed: 9775516.
6. Eerola T, Himberg T, Toiviainen P, Louhivuori J (2006) Perceived complexity of western and African folk melodies by western and African listeners. Psychol Music 34: 337-371. doi:10.1177/0305735606064842.
7. Berlyne DE (1960) Conflict, arousal and curiosity. New York: McGraw-Hill.
8. Berlyne DE (1970) Novelty, complexity and hedonic value. Percept Psychophys 8: 279-286. doi:10.3758/BF03212593.
9. Berlyne DE (1971) Aesthetics and psychobiology. New York: Appleton-Century-Crofts.
10. Berlyne DE (1974) Novelty, complexity, and interestingness In: DE Berlyne. Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic comparison. Washington: Hemisphere. pp. 175-180.
11. Casti J, Karlqvist A (2003) Art and complexity. Amsterdam: Elsevier.
12. Chatterjee A (2003) Prospects for a cognitive neuroscience of visual aesthetics. Bulletin Psych Arts 4: 55-60.
13. Leder H, Belke B, Oeberst A, Augustin D (2004) A model of aesthetic appreciation and aesthetic judgements. Br J Psychol 95: 489-508. doi:10.1348/0007126042369811. PubMed: 15527534.
14. McDermott JH (2012) Auditory preferences and aesthetics: Music, voices, and everyday sounds. In: RJ DolanT Sharot. Neuroscience of preference and choice: Cognitive and neural mechanisms. London: Academic Press. pp. 227-296.
15. Vartanian O, Nadal M (2007) A biological approach to a model of aesthetic experience. In: L DorfmanC MartindaleV Petrov. Aesthetics and innovation. Newcastle: Cambridge Scholars Publishing. pp. 429-444.
16. Nadal M, Munar E, Marty G, Cela-Conde CJ (2010) Visual complexity and beauty appreciation: Explaining the divergence of results. Empirical Stud Arts 28: 173-191. doi:10.2190/EM.28.2.d.
17. Berlyne DE, Ogilvie JC, Parham LCC (1968) The dimensionality of visual complexity, interestingness, and pleasingness. Can J Psychol 22: 376-387. doi:10.1037/h0082777. PubMed: 5724480.
18. Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39: 1161-1178. doi:10.1037/h0077714.
19. Bradley MM, Codispoti M, Sabatinelli D, Lang PJ (2001) Emotion and motivation II: Sex differences in picture processing. Emotion 1: 300-319. doi:10.1037/1528-3542.1.3.300. PubMed: 12934688.
20. Nater UM, Abbruzzese E, Krebs M, Ehlert U (2006) Sex differences in emotional and psychophysiological responses to musical stimuli. Int J Psychophysiol 62: 300-308. doi:10.1016/j.ijpsycho.2006.05.011. PubMed: 16828911.
21. Berlyne DE (1963) Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. Can J Psychol 17: 274-290. doi:10.1037/h0092883. PubMed: 14048839.
22. Munsinger H, Kessen W (1964) Uncertainty, structure, and preference. Psychol Monogr Gen Appl 78: 1-24.
23. Day H (1967) Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. Percept Psychophys 2: 281-286. doi:10.3758/BF03211042.
24. Farley FH, Weinstock CA (1980) Experimental aesthetics: Children's complexity preference in original art and photoreproductions. Bull Psychon Soc 15: 194-196.
25. Imamoglu C (2000) Complexity, liking and familiarity: Architecture and nonarchitecture Turkish students' assessment of traditional and modern house facades. J Environ Psychol 20: 5-16. doi:10.1006/jevp.1999.0155.
26. Saklofske DH (1975) Visual aesthetic complexity, attractiveness and diverse exploration. Percept Mot Skills 41: 813-814. doi:10.2466/pms.1975.41.3.813. PubMed: 1215122.
27. Wohlwill JF (1968) Amount of stimulus exploration and preference as differential functions of stimulus complexity. Percept Psychophys 4: 307-312. doi:10.3758/BF03210521.
28. Beauvois MW (2007) Quantifying aesthetic preference and perceived complexity in fractal melodies. Music Percept 24: 247-264. doi:10.1525/mp.2007.24.3.247.
29. Bragg BWE, Crozier JB (1974) The development with age of verbal and exploratory responses to sound sequences varying in uncertainty level. In: DE Berlyne. Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic comparison. Washington: Hemisphere. pp. 91-108.
30. Burke MJ, Gridley MC (1990) Musical preferences as a function of stimulus complexity and listeners' sophistication. Percept Mot Skills 71: 687-690. doi:10.2466/pms.1990.71.2.687.
31. Crozier BJ (1974) Verbal and exploratory responses to sound sequences varying in uncertainty level. In: DE Berlyne. Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic comparison. Washington: Hemisphere. pp. 27-90.
32. Heyduck RG (1975) Rated preference for musical compositions as it relates to complexity and exposure frequency. Percept Psychophys 17: 84-91. doi:10.3758/BF03204003.
33. North AC, Hargreaves DJ (1995) Subjective complexity, familiarity and liking for popular music. Psychomusicology 14: 77-93. doi:10.1037/h0094090.
34. Steck L, Machotka P (1975) Preference for musical complexity: Effects of context. J Exp Psychol Hum Percept Perform 1: 170-174. doi:10.1037/0096-1523.1.2.170.
35. Vitz PC (1966) Affect as a function of stimulus variation. J Exp Psychol Hum Learn 71: 74-79. PubMed: 5902145.
36. Kreitler S, Zigler E, Kreitler H (1974) The complexity of complexity. Hum Dev 17: 54-73. doi:10.1159/000271333. PubMed: 4833151.
37. Krupinski E, Locher P (1988) Skin conductance and aesthetic evaluative responses to non-representational works of art varying in symmetry. Bull Psychon Soc 26: 355-358.
38. Neperud RW, Marschalek DG (1988) Informational and affect bases of aesthetic response. Leonardo 21: 305-312. doi:10.2307/1578660.
39. Messinger SM (1998) Pleasure and complexity: Berlyne revisited. J Psychol 132: 558-560. doi:10.1080/00223989809599288.
40. Nicki RM, Moss V (1975) Preference for non-representational art as a function of various measures of complexity. Can J Psychol 29: 237-249. doi:10.1037/h0082029.
41. Osborne JW, Farley FH (1970) The relationship between aesthetic preference and visual complexity in abstract art. Psychon Sci 19: 69-70.
42. Eisenberg J, Thompson WF (2003) A matter of taste: Evaluating improvised music. Creativity Res J 25: 287-296.

43. Martindale C, Moore K (1989) Relationship of musical preference to collative, ecological, and psychophysical variables. Music Percept 6: 432-446.
44. Orr MG, Ohlsson S (2005) Relationship between complexity and liking as a function of expertise. Music Percept 22: 583-611. doi:10.1525/mp.2005.22.4.583.
45. Russell PA (1982) Relationships between judgements of the complexity, pleasingness and interestingness of music. Current Psychological Res 2: 195-202. doi:10.1007/BF03186760.
46. Smith JD, Melara RJ (1990) Aesthetic preference and syntactic prototypically in music: Tis the gift to be simple. Cognition 34: 279-298. .
47. Vitz PC (1964) Preferences for rates of information presented by sequences of tones. J Exp Psychol Hum Learn 68: 176-183. PubMed: 14194406.
48. Konečni VJ (1976) Daniel E Berlyne: 1924-1976. Am J Psychol 91: 133-137.
49. Aksentijevic A, Gibson K (2012) Psychological complexity and the cost of information processing. Theory Psychol. doi: 10.1177/0959354311423435.
50. Streich S (2007) Music complexity: A multi-faceted description of audio content. Doctoral dissertation, University of Pompeu Fabra, Barcelona.
51. Martindale C (1984) The pleasures of thought: A theory of cognitive hedonics. J Mind Behav 5: 49-80.
52. Martindale C (1988) Aesthetics, psychobiology, and cognition. In: F FarleyR Neperud. The foundations of aesthetics, art, and art education. Nueva York. Praeger. pp. 7-42
53. Barron F, Welsh GS (1952) Artistic perception as a possible factor in personality style: Its measurement by a figure preference test. J Psychol 33: 199-203. doi:10.1080/00223980.1952.9712830.
54. Eisenman R, Cofee S (1964) Aesthetic preferences of art students and mathematics students. J Psychol 58: 375-378. doi: 10.1080/00223980.1964.9916759.
55. Taylor RE, Eisenman R (1964) Perception and production of complexity by creative art students. J Psychol 57: 239-242. doi: 10.1080/00223980.1964.9916693. PubMed: 14100125.
56. Furnham A, Walker J (2001) The influence of personality traits, previous experience of art and demographic variables on artistic preference. Pers Individ Dif 31: 997-1017. doi:10.1016/S0191-8869(00)00202-6.
57. Parncutt R, Marin MM (2006) Emotions and associations evoked by unfamiliar music. In: H GottesdienerJC Vilatte. Culture and communication: Proceedings of the XIXth Conference of the International Association of Empirical; Aesthetics Laboratoire Culture et Communication, Université d'Avignon et des Pays de Vaucluse. pp. 725-729
58. Schäfer T, Sedlmeier P (2010) What makes us like music? Determinants of music preference. Psych Aesthet Creativity and Arts 4: 223-234. doi:10.1037/a0018374.
59. Barron F (1952) Personality style and perceptual choice. J Pers 20: 385-401. doi:10.1111/j.1467-6494.1952.tb01116.x. PubMed: 14939140.
60. Chamorro-Premuzic T, Burke C, Hsu A, Swami V (2010) Personality predictors of artistic preference as a function of the emotional valence and perceived complexity of paintings. Psych Aesthet Creativity and Arts 4: 196-204. doi:10.1037/a0019211.
61. Rentfrow PJ, Gosling SD (2003) The do re mi's of everyday life: The structure and personality correlates of music preferences. J Pers Soc Psychol 84: 1236-1256. doi:10.1037/0022-3514.84.6.1236. PubMed: 12793587.
62. Feist GJ, Brady TR (2004) Openness to experience, non-conformity, and the preference for abstract art. Empirical Stud Arts 22: 77-89. doi: 10.2190/Y7CA-TBY6-V7LR-76GK.
63. Konečni VJ (2010) The influence of affect on music choice. In: PN JuslinJA Sloboda, Handbook of music and emotion: Theory, research, applications. Oxford: Oxford University Press. pp. 697-723.
64. Cupchik GC (1988) The legacy of Daniel E Berlyne. Empirical Stud Arts 6: 171-186. doi:10.2190/FLM8-6NQ7-N5WM-WLLT.
65. Tan ES (2000) Emotion, art, and the humanities. In: Lewis M & JM Haviland-Jones, JM, editors, Handbook of emotions (2nd ed, pp 116-134). New York: Guilford Press. pp. 116-136
66. Juslin PN, Sloboda JA (2010) Handbook of music and emotion: Theory, research, applications. Oxford: Oxford University Press.
67. Hager M, Hagemann D, Danner D, Schankin A (2012) Assessing aesthetic appreciation of visual artworks – The construction of the Art Reception Survey (ARS). Psych Aesthet Creativity and Arts 6: 320-333. doi:10.1037/a0028776.
68. Hagtvedt H, Hagtvedt R, Patrick VM (2008) The perception and evaluation of visual art. Empirical Stud Arts 26: 197-218. doi: 10.2190/EM.26.2.d.
69. Silvia PJ (2005) Cognitive appraisals and interest in visual art: Exploring an appraisal theory of aesthetic emotions. Empirical Stud Arts 23: 119-133. doi:10.2190/12AV-AH2P-MCEH-289E.
70. Silvia PJ (2005) Emotional responses to art: From collation and arousal to cognition and emotion. Rev Gen Psychol 9: 342-357. doi: 10.1037/1089-2680.9.4.342.
71. Clore GL, Ortony A (2010) Appraisal theories: How cognition shapes affect in emotion. Lewis M HavilandJones JM, L Feldman Barrett L, editors. Handbook of emotions. New York: Guilford Press. pp. 628-642
72. Cupchik GC (1995) Emotions in aesthetics: Reactive and reflective models. Poetics 23: 177-188. doi:10.1016/0304-422X(94)00014-W.
73. Zentner M, Grandjean D, Scherer KR (2008) Emotions evoked by the sound of music: Characterization, classification, and measurement. Emotion 8: 494-521. doi:10.1037/1528-3542.8.4.494. PubMed: 18729581.
74. Forsythe A, Nadal M, Sheehy N, Cela-Conde CJ (2011) Predicting beauty: Fractal dimension and visual complexity in art. Br J Psychol 102: 49-70. doi:10.1348/000712610X498958. PubMed: 21241285.
75. Jacobsen T, Höfel L (2002) Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. Percept Mot Skills 95: 755-766. doi:10.2466/pms.2002.95.3.755. PubMed: 12509172.
76. Keller PE, Schubert E (2011) Cognitive and affective judgements of syncopated musical themes. Adv Cogn Psychol 7: 142-156. doi: 10.2478/v10053-008-0094-0. PubMed: 22253676.
77. Bermpohl F, Pascual-Leone A, Amedi A, Merabet LB, Fregni F et al. (2006) Dissociable networks for the expectancy and perception of emotional stimuli in the human brain. NeuroImage 30: 588-600. doi: 10.1016/j.neuroimage.2005.09.040. PubMed: 16275018.
78. Schupp HT, Junghöfer M, Weike AI, Hamm AO (2003) Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. Neuroreport 14: 1107-1110. doi:10.1097/00001756-200306110-00002. PubMed: 12821791.
79. Scherer KR (1984) On the nature and function of emotion: A component process approach In: KR SchererP Ekman. Approaches to emotion. Hillsdale: Erlbaum. pp. 293-317.
80. Osgood C, Suci G, Tannenbaum P (1957) The measurement of meaning. Urbana, IL: University of Illinois.
81. Berlyne DE (1966) The semantic differential and other measures of reaction to visual complexity. Can J Psychol 20: 125-135. doi:10.1037/h0082931. PubMed: 5942316.
82. Koole S (2009) The psychology of emotion regulation: An integrative review. Cogn Emotion 23: 4-41. doi:10.1080/02699930802619031.
83. Marin MM, Gingras B, Bhattacharya J (2012) Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. Emotion 12: 618-631. doi:10.1037/a0025020. PubMed: 21859191.
84. Brody LR, Hall JA (2009) Gender and emotion in context. In: M LewisJM Haviland-JonesL Feldman Barrett. Handbook of emotions. New York: Guilford Press. pp. 395-408.
85. Cela-Conde CJ, Ayala FJ, Munar E, Maestú F, Nadal M et al. (2009) Sex-related similarities and differences in the neural correlates of beauty. Proc Natl Acad Sci U S A 106: 3847-3852. doi:10.1073/pnas.0900304106. PubMed: 19237562.
86. Cupchik GC, Gebotys R (1988) The experience of time, pleasure, and interest during aesthetic episodes. Empirical Stud Arts 6: 1-12. doi: 10.2190/5YN3-J3P8-FWHY-UDB3.
87. Gard MG, Kring AM (2007) Sex differences in the time course of emotion. Emotion 7: 429-437. doi:10.1037/1528-3542.7.2.429. PubMed: 17516819.
88. Sharp C, van Goozen S, Goodyer I (2006) Children's subjective emotional reactivity to affective pictures: Gender differences and their antisocial correlates in an unselected sample of 7-11-year-olds. J Child Psych And Allied Discip 47: 143-150. doi:10.1111/j.1469-7610.2005.01464.x.
89. Mielby LH, Kildegaard H, Gabrielsen G, Edelenbos M, Thybo AK (2012) Adolescent and adult visual preferences for pictures of fruit and vegetable mixes – effects of complexity. Food Qual Preference 26: 188-195. doi:10.1016/j.foodqual.2012.04.014.
90. Jellinek JS, Köster EP (1979) Perceived fragrance complexity and its relation to familiarity and pleasantness. J Soc Cosmet Chem 30: 253–262.
91. Jellinek JS, Köster EP (1983) Perceived fragrance complexity and its relation to familiarity and pleasantnessII. J Soc Cosmet Chem 34: 83–97.
92. Smith JK, Smith LF (2001) Spending time on art. Empirical Stud Arts 19: 229-236. doi:10.2190/5MQM-59JH-X21R-JN5J.

93. Eerola T (2011) Are the emotions expressed in music genre-specific? An audio-based evaluation of datasets spanning classical, film, pop and mixed genres. J New Musical Res 40: 349-366. doi: 10.1080/09298215.2011.602195.
94. Forsythe A (2009) Visual complexity: Is there all there is? In: D Harris. Engineering psychology and cognitive ergonomics. Berlin: Springer Verlag. pp. 158-166.
95. Gartus A, Leder H (2011) Predicting the visual complexity of line drawings and photos of common objects. Perception 40: ECVP Abstract Supplement 206
96. Birkhoff GD (1933) Aesthetic measure Cambridge. Harvard University Press.
97. Eysenck HJ (1941) The empirical determination of an aesthetic formula. Psychol Rev 48: 83-92. doi:10.1037/h0062483.
98. Eysenck HJ, Castle M (1970) Training in art as a factor in the determination of preference judgments for polygons. Br J Psychol 61: 65-81. doi:10.1111/j.2044-8295.1970.tb02802.x. PubMed: 5414910.
99. Bates E, D'Amico S, Jacobsen T, Székely A, Andonova E et al. (2003) Timed picture naming in seven languages. Psychon Bull Rev 10: 344-380. doi:10.3758/BF03196494. PubMed: 12921412.
100. Chikhman V, Bondarko V, Danilova M, Goluzina A, Shelepin Y (2012) Complexity of images: Experimental and computational estimates compared. Perception 41: 631-647. doi:10.1068/p6987. PubMed: 23094454.
101. Donderi DC (2006) An information theory analysis of visual complexity and dissimilarity. Perception 35: 823-835. doi:10.1068/p5249. PubMed: 16836047.
102. Forsythe A, Mulhern G, Sawey M (2008) Confounds in pictorial sets: The role of complexity and familiarity in basic level pictures processing. Behav Res Methods 40: 116-129. doi:10.3758/BRM.40.1.116. PubMed: 18411534.
103. Forsythe A, Sheehy N, Sawey M (2003) Measuring icon complexity: An automated analysis. Behav Res Methods Instrum Comput 35: 334-342. doi:10.3758/BF03202562. PubMed: 12834094.
104. Cavalcante A, Mansouri A, Mansouri LKE, Kardec B, Takeuchi Y et al. (2011) On the relation between contrast and the perception of visual complexity. Proceedings of the 21st Annual Conference of the Japanese Neural Network Society December. p. 2011). Available: http://jnnsorg/conference/misc/camera_ready/P2-22pdf. Accessed 15 July 2012
105. Donderi DC, McFadden S (2005) Compressed file length predicts search time and errors on visual displays. Displays 26: 71-78. doi: 10.1016/j.displa.2005.02.002.
106. Rossion B, Pourtois G (2005) Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface information in basic level object recognition. Perception 33: 217-236.
107. Snodgrass JG, Vanderwart M (1980) A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. J Exp Psychol Hum Learn Mem 6: 174-215. doi: 10.1037/0278-7393.6.2.174.
108. Li M, Vitanyi P (1997) An introduction to Kolmogorov complexity and its applications. New York: Springer-Verlag.
109. Solomonoff RJ (1986) The application of algorithmic probability to problems in artificial intelligence. In: LN KanalJF Lemmer. Uncertainty in artificial intelligence. North-Holland: Elsevier. pp. 473-491
110. Boolos G, Burgess J, Jeffrey R (2002) Computability and logic. Cambridge: Cambridge University Press.
111. Zhang D, Lu G (2004) Review of shape representation and description techniques. Pattern Recognit 37: 1-19. doi:10.1016/j.patcog. 2003.07.008.
112. Boon JP, Casti J, Taylor RP (2011) Artistic forms and complexity nonlinear dynamics. Psych Life Sciences 15: 265-283.
113. Pearce MT, Wiggins G (2006) Expectation in melody: The influence of context and learning. Music Percept 23: 377-405. doi:10.1525/mp. 2006.23.5.377.
114. Pearce MT, Ruiz MH, Kapasi S, Wiggins GA, Bhattacharya J (2010) Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation. NeuroImage 50: 302-313. doi:10.1016/j.neuroimage.2009.12.019. PubMed: 20005297.
115. Mauch M, Levy M (2011) Structural change on multiple time scales as a correlate of musical complexity. Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011). Available. http://matthiasmauch.net/_pdf/mauch_scm_2011.pdf. Accessed 15 April 2013.
116. Cilibrase R, Vitány P, de Wolf R (2004) Algorithmic clustering of music based on string compression. Comput Music J 28: 49-67. doi: 10.1162/014892604323112248.
117. Londei A, Loreto V, Belardinelli MO (2003) Musical style and authorship categorization by informative compressors. In: R KopiezAC LehmannI WoltherC Wolf. . In: Proceedings of the 5th Triennial ESCOM Conference Liege. Belgium: European Society for the Cognitive Sciences of Music. pp 200-203.
118. Vettehen PH, Nuijten K, Peeters A (2008) Explaining effects of sensationalism on liking of television news stories: The role of emotional arousal. Commun Res 35: 319-338. doi: 10.1177/0093650208315960.
119. Fernandez D, Wilkins AJ (2008) Uncomfortable images in art and nature. Perception 37: 1098-1113. doi:10.1068/p5814. PubMed: 18773732.
120. Graham DJ, Field DJ (2007) Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. Spat Vis 21: 149-164. doi:10.1163/156856807782753877. PubMed: 18073056.
121. Redies C, Hasenstein J, Denzler J (2007) Fractal-like image statistics in visual art: Similarity to natural scenes. Spat Vis 21: 137-148. doi: 10.1163/156856807782753921. PubMed: 18073055.
122. Lang PJ, Bradley MM, Cuthbert BN (2005) International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report p. A-7 Gainesville, FL: University of Florida.
123. Kovesi P (2003) Phase congruency detects corners and edges. Sun C, Taöbot H, Ourselin S, Adriaansen T, editors. Proceedings of the VIIth Biennial Australian Pattern Recognition Society Conference: DICTA 2003. Collingwood: CSIRO. pp 309-318
124. Gonzalez RC, Woods RE, Eddins SL (2003) igital image processing using MATLAB. New Jersey: Prentice Hall.
125. Lartillot O, Toiviainen P, Eerola T (2007) A Matlab toolbox for music information retrieval. In: C PreisachH BurkhardtL Schmidt-ThiemeR Decker. Data analysis, machine learning and applications. Studies in classification, data analysis, and knowledge-organization. Berlin: Springer Verlag. pp. 261-268.
126. Bradley MM, Hamby S, Löw A, Lang PJ (2007) Brain potentials in perception. Picture complexity and emotional arousal. Psychophysiology 44: 364-373. doi:10.1111/j.1469-8986.2007.00520.x. PubMed: 17433095.
127. Gabrielsson A (2002) Emotion perceived and emotion felt: Same or different? Musicae Sci special issue: 2001–2002: 123-147
128. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9: 62-66. doi:10.1109/TSMC. 1979.4310076.
129. Canny J (1986) A computational approach to edge detection. IEEE. Transactions on Pattern Analysis and Machine Intelligence, PAMI-8: 679-698.
130. Peli E (1990) Contrast in complex images. J Opt Soc Am A 7: 2032-2040. doi:10.1364/JOSAA.7.002032. PubMed: 2231113.
131. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Statist 6: 65-70.
132. Hayes AF, Preacher KJ (2010) Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. Multivariate Behav Res 45: 627-660. doi: 10.1080/00273171.2010.498290.
133. Shrout PE, Fleiss JL (1979) Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 2: 420-428 PubMed: 18839484.
134. McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. Psychol Methods 1: 30-46. doi: 10.1037/1082-989X.1.1.30.
135. Rozin P, Royzman EB (2001) Negativity bias, negativity dominance, and contagion. Pers Soc Psychol Rev 5: 296-320. doi:10.1207/ S15327957PSPR0504_2.
136. Myers L, Sirois MJ (2006) Spearman correlation coefficients, differences between. Encyclopedia of Statistical Sciences John Wiley & Sons, Inc. Available: DOI: 101002/ 0471667196ess5050pub2. Accessed 15 July 2012
137. Hayes AF (2009) Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. Commun Monogr 76: 408-420. doi: 10.1080/03637750903310360.
138. Rosenthal R (1991) Meta-analytic procedures for social research. Newbury Park, CA: Sage.
139. Lin Y-C, I B (2011) Generation and gender differences in aesthetic responses to popular illustrations. Visual Arts Research 37: 30-41
140. Meyers-Levy J, Zhu R (2010) Gender differences in the meanings consumers infer from music and other aesthetic stimuli. J Consum Psychol 20: 495-507. doi:10.1016/j.jcps.2010.06.006.
141. Pandey S, Elliott W (2010) Suppressor variables in social work research: Ways to identify in multiple regression models. J Soc Soc Works Res 1: 28-40. doi:10.5243/jsswr.2010.2.
142. Thayer RE (1989) The biopsychology of mood and arousal. New York: Oxford University Press.
143. Kettunen J, Ravaja N, Näätänen P, Keskivaara P, Keltikangas-Järvinen L (1998) The synchronization of electrodermal activity and its

relationship to energetic arousal: A time series approach. Biol Psychol 48: 209-225. doi:10.1016/S0301-0511(98)00017-9. PubMed: 9788761.
144. Kreibig SD (2012) Emotion, motivation, and cardiovascular response. In: RA WrightGHE Gendolla. Motivation on cardiovascular response. Washington: American Psychological Association. pp. 93-117.
145. Lacey JI, Kagan J, Lacey BC, Moss HA (1963) The visceral level: Situational determinants and behavioral correlates of autonomic response patterns. In: PH Knapp. Expressions of the emotions in man New York: International Universities Press. pp. 161-196
146. Mauss IB, Levenson RW, McCarter L, Wilhelm FH, Gross JJ (2005) The tie that binds? Coherence among emotion experience, behavior, and physiology. Emotion 5: 175-190. doi:10.1037/1528-3542.5.2.175. PubMed: 15982083.
147. Clore GL, Huntsinger JR (2007) How emotions inform judgements and regulate thought. Trends Cogn Sci 11: 393-399. doi:10.1016/j.tics.2007.08.005. PubMed: 17698405.
148. Storbeck J, Clore GL (2008) Affective arousal as information: How affective arousal influences judgments, learning and memory. Soc Pers Psych Compass 2: 1824-1843. doi:10.1111/j.1751-9004.2008.00138.x.
149. Martindale C, Moore K, Borkum J (1990) Aesthetic preference: Anomalous findings for Berlyne's psychobiological model. Am J Psychol 103: 53-80. doi:10.2307/1423259.
150. Leder H, Gerger G, Dressler SG, Schabmann A (2012) How art is appreciated. Psych Aesthet Creativity and Arts 6: 2-10. doi:10.1037/a0026396.
151. Rentfrow PJ, Goldberg LR, Levitin DJ (2011) The structure of musical preference: A five-factor model. J Pers Soc Psychol 100: 1139-1157. doi:10.1037/a0022406. PubMed: 21299309.
152. Schäfer T, Sedlmeier P (2011) Does the body move the soul? The impact of arousal on music preference. Music Percept 29: 37-50. doi:10.1525/mp.2011.29.1.37.
153. Kovesi P (1999) Image features from phase congruency. Videre J Computer Vision Res 1: 2-26.
154. Trochidis K, Delbé C, Bigand E (2011) Investigation of the relationship between audio features and induced emotions in contemporary western music. Proceedings of the 8th Sound and Music Computing Conference 2011. Available. http://smcnetworkorg/smc_papers-2011114. Accessed 15 July 2012.
155. Alais D, Newell FN, Mamassian P (2010) Multisensory processing in review: From physiology to behavior. Seeing and Perceiving 23: 3-38.
156. Shimojo S, Shams L (2001) Sensory modalities are not separate modalities: Plasticity and interactions. Curr Opin Neurobiol 11: 505-509. doi:10.1016/S0959-4388(00)00241-5. PubMed: 11502399.
157. Baumgartner T, Lutz K, Schmidt KF, Jäncke L (2006) The emotional power of music: How music enhances the feeling of affective pictures. Brain Res 1075: 151-164. doi:10.1016/j.brainres.2005.12.065. PubMed: 16458860.
158. Logeswaran N, Bhattacharya J (2009) Crossmodal transfer of emotion by music. Neurosci Lett 455: 129-133. doi:10.1016/j.neulet.2009.03.044. PubMed: 19368861.
159. Lithari C, Frantzidis CA, Papadelis C, Vivas AB, Klados MA et al. (2010) Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. Brain Topogr 23: 27-40. doi:10.1007/s10548-009-0130-5. PubMed: 20043199.
160. Kreutz G, Schubert E, Mitchell L (2008) Cognitive styles of music listening. Music Percept 26: 57-73. doi:10.1525/mp.2008.26.1.57.
161. Silverman I, Eals M (1992) Sex differences in spatial abilities: Evolutionary theory and data. In: J BarkowL CosmidesJ Tooby. The adapted mind: Evolutionary psychology and the generation of culture. New York: Oxford University Press. pp 487-503.
162. Abramov I, Gordon J, Feldman O, Chavarga A (2012). Sex Vision II: color appearance of monochromatic lights. Biology of Sex Differences 3: 21. Available: http://www.bsd-journal.com/content/3/1/21. Accessed 10 September 2012

2.1.4 Summary and outlook

This study comprised four experiments, which were similar in terms of the procedure and analyses employed, but which differed in terms of the stimulus categories. The data analysis, for which the unit of analysis was the picture and not the rater, was conducted for males and females separately and also on combined ratings. In Experiment 1, participants rated environmental scenes. Notably, familiarity emerged as an important variable that should be controlled for in statistical analyses (e.g., in correlations between other subjective variables and regression analyses). Familiarity was negatively correlated with complexity and arousal, and positively with pleasantness, with the largest effect on pleasantness ratings. The pleasantness-arousal emotion space covered by our stimuli was U-shaped and similar to previous reports involving IAPS pictures. Importantly, we observed a positive relationship between subjective complexity and arousal (r_s = .27 controlled for familiarity) and no relationship between complexity and pleasantness when ratings were averaged per picture. Arousal was identified as a significant mediator in the complexity-pleasantness relationship by running a mediated regression analysis (55% of the variance in pleasantness was explained when controlling familiarity). There was no direct effect of complexity on pleasantness, but "[at] all levels of picture complexity, an increase of complexity led to a decrease in pleasantness through the effect of complexity on arousal" (Marin & Leder, 2013, p. 9). These results were similar in males and females and even held true when familiarity was not considered in the model. The mean RMS contrast measure turned out to be the most efficient predictor (β = .55, p < .001) of subjective complexity in a regression analysis controlling for familiarity and considering arousal, pleasantness and the objective measures of complexity. Compression formats such as JPEG (r_s = .52) and TIFF (r_s = .53) worked moderately well when considering zero-order correlations. A regression model only considering objective measures revealed that the mean RMS contrast measure was the only significant predictor (β = .54, p < .001).

The results of Experiment 2, in a nutshell, were that the emotion space spanned by representational paintings was similar in shape to the one covered by environmental scenes. However, the stimuli were less dispersed. Familiarity also correlated positively with pleasantness and negatively with arousal, but not with complexity due to gender differences (negative correlation among females, positive correlation for males). Notably, the association between complexity and arousal was sig-

nificantly stronger for paintings (r_s = .59) than for environmental scenes (r_s = .27). There was no clear relationship between complexity and pleasantness as was the case in Experiment 1. There were several gender differences when analyzing the correlations between the subjective measures, which were not present in Experiment 1. Significant indirect effects of arousal on pleasantness only emerged when the genders were analyzed separately, or when familiarity was not considered in the model. Our set of objective complexity measures did not work as well for paintings as for environmental scenes: only the standard deviation of the phase congruency edge detection measure correlated significantly with subjective complexity (r_s = .38), but none of the compressed file-size measures yielded significant zero-order correlations. Interestingly, some of the objective complexity measures also correlated with subjective arousal. A multiple regression model considering both subjective and objective measures and controlling for familiarity showed that arousal (β = .56, p < .001) was a better predictor of complexity than the standard deviation phase congruency measure (β = .28, p = .013). A regression model only considering objective measures as predictors yielded significant results for the *SD* measure of phase congruency (β = .69, p <.001), the GIF compression file size (β = .39, p = .003) and the *SD* measure of RMS contrast (β = -.30, p = .039). Only 25% of the variance in subjective complexity could be explained, whereas a similar model for environmental scenes explained 28% of the variance with only one predictor, namely the mean RMS contrast measure.

These two experiments provided preliminary evidence and theoretical support for Berlyne's (1971) psychobiological model in the visual domain by studying affective stimuli with semantic contents. The importance of arousal was demonstrated by a positive relationship with complexity (even when controlling for familiarity), which was stronger in representational paintings compared to environmental scenes. The arousal-complexity relationship in responses to environmental scenes has been studied by Madan et al. (2018), who argued that a small proportion of the arousal-complexity bias may be related to semantic associations evoked by the environmental scenes and that it is yet unclear why arousing pictures are more complex. In any case, whether complexity leads to an increase of arousal, or whether arousal leads to an increase of perceived complexity, cannot be disentangled in a correlational study. Moreover, the existence of a mediation effect of arousal in the complexity-pleasantness relationship was shown as predicted by Berlyne (1971). However,

probably due to how the stimuli were pre-selected, there were no indications of an inverted U-shaped relationship between complexity and pleasantness. In the case of representational paintings, we detected several non-negligible gender effects, which highlight the need to consider gender more appropriately in future studies (see, e.g., Madan et al., 2018, for a similar argument when studying affective environmental scenes).

Another key finding was that objective measures of complexity performed differently for the two stimulus sets, which can partly be explained by the presence of brush strokes in the case of paintings. Compression size worked better for environmental scenes than for paintings, and objective complexity of the two stimulus sets was predicted by different types of edge detection algorithms, although RMS contrast measures seem to be well suited for both types of stimuli. In environmental scenes, the objective complexity measures did not correlate with arousal ratings, corroborating findings by Madan et al. (2018). In Experiment 2, however, we observed that arousal correlated with some complexity measures of representational paintings, which shows that affect and visual complexity may interact more strongly in art perception than everyday scene perception (see also Experiments 3 and 4). Although our studies comprised 13 measures of objective complexity based on edge detection and compressed file size and showed that linear combinations of these perform differently for our two stimulus sets, one may nowadays think of including other complexity measures (Braun et al., 2013; Carballal et al., 2018; Carballal et al., 2020; Durmus, 2020; Machado et al., 2015; Madan et al., 2018; Nagle & Lavie, 2020; Nath et al., 2024; Redies et al., 2012; Sun et al., 2015; Zhou et al., 2023), more advanced statistical data analysis methods (Fernandez-Lozano et al., 2019), such as random forest models (Gartus & Leder, 2017), as well as much larger stimulus sets (Braun et al., 2013; Karjus et al., 2023) and participant groups (Zhou et al., 2023). Future research may also include objective measures that try to capture the fractal nature of visual scenes (Forsythe et al., 2011; Friedenberg et al., 2022; Robles et al., 2021; Taylor et al., 2005), and we strongly suggest employing linear combinations of several types of objective measures to capture visual complexity instead of a single one. It should also be possible to integrate image emotion analysis (Zhu et al., 2024) with complexity analysis in the future.

Forsythe et al. (2017) studied the relationship between subjective complexity, familiarity, beauty and compression file measures using the Rossion and Pourtois

image set (2004) for colorized, grayscale and line-drawings of simple objects. They also reported a negative relationship between complexity and familiarity, which thus seems to be present when studying various visual stimulus sets, and suggest that this needs to be taken into account when developing normative ratings of image sets. Moreover, Forsythe et al. (2017) found, as we did, that familiarity does not correlate with automated measures and suggest that objective measures of complexity are thus a better, unbiased measure than subjective complexity. They also studied two measures related to familiarity, namely familiarity with the image and previous encounter with the object, and both correlated highly with each other. This is an interesting finding with regard to how we studied familiarity: we asked for familiarity ratings referring to the contents of the image for environmental scenes in Experiment 1, and familiarity with the painting in Experiment 2 (which is more closely to encounter in Forsythe et al., 2017). Future studies may thus study familiarity for scenes with and without semantic contents more closely, with several different measures related to familiarity to find out which is the most appropriate and efficient in this context. In short, Forsythe et al. (2017) also conclude that familiarity is an important measure in empirical aesthetics when studying complexity (see also Forsythe et al., 2008; Snodgrass & Vanderwart, 1980).

In Experiment 3, following a similar approach as in Experiments 1 and 2, the emotion space covered by the musical excerpts did not show a U-shape as was the case for our visual stimuli. Familiarity correlated positively with pleasantness and negatively, but not significantly, with arousal. Due to a gender effect, complexity did not correlate with familiarity when males and females were considered together, which we also observed for Experiment 2. Nevertheless, this suggests that familiarity acts similarly, but to various degrees, on arousal and pleasantness across sensory and artistic domains, with the largest effects on pleasantness followed by arousal. In general, familiarity affected environmental scene ratings to a larger degree than those of paintings and music. The complexity-arousal relationship differed significantly across the three stimulus sets, and it was the strongest for music (r_s = .82), followed by paintings (r_s = .59) and environmental scenes (r_s = .27). Arousal emerged as a mediator in the complexity-pleasantness relationship (mostly driven by females' ratings), but there was also a direct effect of complexity on pleasantness. Direct effects were not observed for visual stimuli, which suggests a critical difference between the sensory domains. All three mediation models had a similar R^2 of around

.55, which is another interesting finding. The four new objective measures of musical complexity, inspired by those employed to the study of visual complexity, all significantly correlated with subjective complexity, and also to a similar degree with arousal. FLAC compression yielded the strongest correlation (r_s = .65). A multiple regression based on subjective and objective measures found that arousal was the strongest predictor of subjective complexity (which was also the case for paintings), followed by pleasantness (β = .28, p = .001) and FLAC file size (β = .20, p = .03). Interestingly, a gender wise analysis yielded a different result and event density and MP3 file size were better predictors than FLAC file size, probably due to differential familiarity effects. Considering only objective measures in the model, it was found that a linear combination of FLAC file size, MP3 file size and event density were successful predictors, explaining around 60% of the variance. It should be noted that this simple model explained subjective musical complexity much better than the respective models for environmental scenes and representational paintings (25-28%).

In Experiment 4, the complexity of the musical stimuli was altered by mixing two musical styles differing in the number of instruments, namely solo piano music and classical piano trios (i.e., piano, violin and cello). Ratings confirmed that solo piano music was perceived as less complex than piano trio music. The emotion space covered by the stimuli was less dispersed than the one for solo piano music. Familiarity correlated positively with pleasantness (moderate to large effects) and negatively with arousal and complexity, which confirms the consistent pattern across the four experiments and suggests that familiarity is an important variable to be considered in related future studies. Complexity correlated significantly with arousal (r_s = .84), as observed in Experiment 3. In addition, the mediation analysis demonstrated that there was a direct effect of complexity on pleasantness, and a significant indirect effect of arousal on pleasantness only in females (i.e., pleasantness decreases when complexity increases through its effect on arousal). These findings basically replicate those reported in Experiment 2 for solo piano music. Future research will have to clarify whether these results hold true for non-classical musical styles. Regarding the objective complexity measures, again the FLAC file size (r_s = .78) was the best predictor of subjective complexity. Event density (r_s = .43) performed worse than in Experiment 3, possibly due to the algorithm performing worse when several voices (instruments) occur in an auditory scene. As in Experiment 3, the objective measures of complexity also correlated with arousal, and in the group of females, also with pleas-

antness. Regression analyses showed that arousal (β = .68, $p < .001$), event density (β = .24, $p < .001$) and MP3 file size (β = .17, $p = .024$) were significant predictors of complexity. The results were similar when both genders were analyzed separately. Finally, a model only comprising objective measures yielded FLAC file size, MP3 file size and event density as significant predictors, explaining around 67% of the variance in musical complexity. Note that this proportion of explained variance in musical complexity is somewhat higher than the one obtained by Streich (2007) with 12 predictors.

Three findings of Experiments 3 and 4 are particularly noteworthy: first, the relationship between subjective complexity and arousal is much stronger than the one observed for the visual domain in general (and note that the one of environmental scenes was the lowest). There were no considerable differences in inter-rater reliability for complexity and arousal ratings across the experiments, which makes it unlikely that participants varied much in their responses and/or experienced, for example, more difficulties to give ratings in one specific experiment. Moreover, we did not observe gender differences when analyzing the complexity-arousal relationship, but gender played a role in the complexity-pleasantness relationship. This suggests that individual differences may play a larger role when measures of hedonic tone are involved in comparison to structural and internal state measures. Second, linear combinations of objective complexity measures performed better for music than for visual stimuli and explained a much larger proportion of the variance in complexity (60-67%). It is thus possible that complexity ratings in music are mostly based on the auditory signal itself, which differs from the visual domain where semantic meaning comes also into play and may generate arousal alongside the contribution of structural aspects of the visual scene (Madan et al., 2018). This hypothesis could easily be tested by obtaining ratings of abstract paintings, also chosen to vary in arousal and pleasantness, and to examine the complexity-arousal relationship in such a stimulus set, which should be stronger than the one observed for representational paintings. The idea that semantic meaning plays a role when explaining the predictive power of objective complexity measures also gains support from recent studies involving natural scenes (Kyle-Davidson et al., 2023, Zhou et al., 2023). By using large stimulus sets and participant groups as well as advanced statistical methods, it has been shown that subjective complexity depends both on structural visual features as well as semantic contents. Moreover, objective complexity could predict subjective

complexity better for scene images than images containing objects (Zhou et al., 2023). Thus, future studies should study different semantic categories in greater depth and consider different subdimensions of visual complexity. This approach could be easily expanded to representational paintings and other visual aesthetic objects. Third, arousal (and sometimes pleasantness) also correlated with objective measures of musical complexity. One may thus speculate that the concepts of arousal and complexity are less clearly differentiated in the musical than in the visual domain, which may also explain why there was an additional direct effect of complexity on pleasantness (besides a mediation effect of arousal) for responses to musical but not to visual stimuli. Lastly, having a look at all four experiments, one notices that objective measures of complexity are more or less unrelated to affective ratings in environmental scenes, partly related to arousal in representational paintings, and most strongly related to musical affect. This may hint at a notable difference between aesthetic and non-aesthetic objects, and at differences between aesthetic objects, which may eventually also be supported by neuroimaging evidence.

In general, our four experiments have shown that familiarity and gender are important variables that should not be neglected in empirical aesthetics, even when being mostly interested in other stimulus-related or person-related variables. For instance, the correlation between familiarity and pleasantness was strong across all four experiments (around r_s = .60). Novelty-familiarity and its effect on aesthetic preference has been discussed within Berlyne's theory (Marin, 2022; Song et al., 2021; Verhaeghen, 2018), but the literature has discussed familiarity without relating to Berlyne as well (Freitas et al., 2018), which points to the relevance of this variable. For example, Zajonc's (1968) *mere-exposure effect* addresses the issue of how repeated exposure to a stimulus changes preference. Conversely, gender is a variable which appears not to be of great interest in the field. However, our experiments have demonstrated that subtle differences between males and females may be present and that gender is a variable that should be studied more closely in empirical aesthetics (see also Coursaris et al., 2008; Johnson & Knapp, 1963; Miller & Hübner, 2023; Tuch et al., 2010).

Altogether, the results of this comprehensive study have provided a solid basis for three other empirical studies included in this doctoral thesis. The follow-up studies focused on the effect of presentation duration (Marin & Leder, 2016), the multidimensionality of hedonic tone (Marin et al., 2016), and the role of individual differences

(Marin & Leder, 2018) when rating the same stimulus material, with the goal of pursuing a systematic comparative approach within the context of Berlyne's model of collative motivation (1971).

2.2 Crossing boundaries: Toward a general model of neuroaesthetics (Marin, 2015)

2.2.1 Introduction and summary

Inspired by the findings of Marin and Leder (2013), this opinion article argues for a comparative approach in the fields of empirical aesthetics and neuroaesthetics. Starting off by giving a short overview of the historical developments of neuroaesthetics with its strong roots in visual aesthetics, this article presents different avenues of how theory building may profit from a comparative approach and the inclusion of a wide range of art forms and other types of objects evoking aesthetic experiences. It is argued that only a systematic comparison of aesthetic responses to different man-made and natural stimulus categories *within* and *across* sensory domains will reveal the characteristics of aesthetic experiences and their underlying brain mechanisms. The proposed research program (which basically reflects the ethos of this doctoral thesis) involves both controlled psychological and neuroscientific experiments, which will ultimately reveal whether broad or narrow theories of aesthetic experiences are more appropriate.

2.2.2 Authorship contribution statement

MMM wrote the article.

2.2.3 Re-print of the published theoretical article

frontiers in Human Neuroscience

OPINION
published: 07 August 2015
doi: 10.3389/fnhum.2015.00443

Crossing boundaries: toward a general model of neuroaesthetics

Manuela M. Marin *

Department of Basic Psychological Research and Research Methods, University of Vienna, Vienna, Austria

Keywords: neuroaesthetics, crossmodal, multimodal, interdisciplinarity, humanities

Since its dawn in the early 2000s, neuroaesthetics has been flowering as an independent research field (Nadal and Skov, 2013). Its emergence has mostly been driven by researchers who specialize in the study of visual perception and cognition, and who show interest in visual arts. This strong link between neuroaesthetics and vision science is reflected not only by the fact that the term "neuroaesthetics" was coined by the renowned vision researcher Semir Zeki (Zeki and Nash, 1999), but also by recent reviews on neuroaesthetics focusing primarily on aesthetic experiences induced by paintings, abstract patterns, landscapes, faces, architecture, fashion and design objects (Cinzia and Vittorio, 2009; Cela-Conde et al., 2011; Chatterjee, 2011, 2014; Nadal, 2013; Chatterjee and Vartanian, 2014). Current models of aesthetic experiences and their brain correlates revolve around the visual modality and include moderators of aesthetic experiences (social, cultural and situational context, personality, expertise etc.) (Ramachandran and Hirstein, 1999; Chatterjee, 2004; Tinio, 2013; Leder and Nadal, 2014; Redies, 2015). These models are not exclusively concerned with the study of beauty or preference, although these concepts are of historical importance, but include a wide range of aesthetic emotions, judgements, and behaviors.

What may be the underlying reasons for this prevalence of studying the neural underpinnings of visual aesthetic experiences? As in any developing research field, it takes a critical number of researchers who are motivated to organize and attend conferences on a given topic (the first conference on Neuroaesthetics took place in 2009, see Nadal and Pearce, 2011), to write books (Zeki and Nash, 1999; Shimamura and Palmer, 2012; Chatterjee, 2013) and to edit special issues in research journals (Nadal and Skov, 2013). Since the vision science community is relatively large in comparison to other research communities in the field of experimental psychology, it is not unanticipated that several prominent figures were able to attract the attention of younger researchers to help launch this fascinating research field. Another reason for the leading status of visual neuroaesthetics may be that, before the advent of neuroscience, interest in empirical aesthetics was stronger in the vision science community than in other research communities. Therefore, the move toward applying neuroimaging methods to questions related to aesthetic experiences seemed the next logical step. Furthermore, the investigation of aesthetic responses to static images of different kinds may be easier than the study of dynamic images, such as film, video art and the performing arts, and the study of dynamic artforms, such as music and poetry.

Lately, the prolific efforts of the visual neuroaesthetics community have attracted the attention of researchers working outside vision science. For example, the number of publications on music, aesthetics and the brain has clearly increased in the last few years (Brattico and Pearce, 2013; Zatorre and Salimpoor, 2013; Koelsch, 2014b), and a neurocognitive model of the aesthetic experience of music was recently published (Brattico et al., 2013). Similarly, research on the neuroaesthetics of literature reception (Bohrn et al., 2013; Jacobs, 2015) and dance (Calvo-Merino et al., 2008; Cross and Ticini, 2012; Christensen and Calvo-Merino, 2013) has opened new pathways to a deeper understanding of aesthetic experiences. With this in mind, it is very refreshing to see that, for instance, comparative brain imaging studies (Ishizu and Zeki, 2011), metanalyses incorporating different perceptual senses (Brown et al., 2011) and reviews considering different artforms (Nadal, 2013) have been published. Furthermore, an interdisciplinary neuroscience of aesthetic experiences

OPEN ACCESS

Edited by:
Oshin Vartanian,
Toronto Research Centre, Canada

Reviewed by:
Zaira Cattaneo,
University of Milano-Bicocca, Italy

***Correspondence:**
Manuela M. Marin,
manuela.marin@univie.ac.at

Received: *15 June 2015*
Accepted: *21 July 2015*
Published: *07 August 2015*

Citation:
Marin MM (2015) Crossing boundaries: toward a general model of neuroaesthetics. Front. Hum. Neurosci. 9:443. doi: 10.3389/fnhum.2015.00443

across the three Sister Arts (paintings, poetry, and music) has recently been proposed (Starr, 2013). Undoubtedly, these pioneering attempts may indicate that neuroaesthetics will move away from a field primarily concerned with visual artforms to a field encompassing different artistic modalities in the future. But can this expansion to other artforms be the ultimate goal of the field?

Although the study of aesthetic experiences across different artforms appears a promising avenue because it may lead to broader, overarching theories of aesthetic experiences, such an endeavor may still remain restrictive. The reasons for this are manifold. A neuroaesthetics of the arts is not in line with the initial, broad definition of aesthetics (*scientia cognitionis sensitivae*) by Baumgarten (2007, 1750/1758), but rather based on the notion that aesthetics concerns the study of the arts and beauty. Moreover, Kant already considered human-made art as well as nature in his reflections on aesthetics in his *Critique of Judgment* (1790), thus suggesting a comprehensive approach to aesthetics. In the last third of the twentieth century, philosophy witnessed a revival of this view by the emergence of environmental aesthetics (Carlson, 2000; Budd, 2002). Purely art-based research in empirical aesthetics is also not in line with the influential empirical approach proposed by Berlyne (1960, 1971) in his "new experimental aesthetics," which conceived aesthetics within the realm of motivation, curiosity, and exploratory behavior in humans and animals. Therefore, a narrow, arts-based approach to neuroaesthetics does not do justice to the frequent induction of aesthetic experiences by objects other than artworks (i.e., natural landscapes, food, faces, bodies, sounds etc.) (Jacobsen, 2006; Brown and Dissanayake, 2009; Tinio, 2013; Zaidel, 2015), but probably also hinders any substantial theoretical developments in the context of evolutionary aesthetics (Grammer et al., 2003). Hence, the neuroaesthetics community will need to follow a comprehensive, broad approach by incorporating the widest possible range of objects unless neuroartsology (Brown and Dissanayake, 2009), i.e., a neuroaesthetics on human art, is in the center of interest. However, this does not entail that neuroaesthetics should become a discipline in which the motto "anything goes" is prevailing. On the contrary, objects of interest should be primarily those for which it has already been shown that they have the capacity to induce aesthetic experiences in humans.

Given that empirical aesthetics in general and neuroaesthetics in particular are lacking a broad theory of aesthetic experiences, which may partly be a byproduct of the general specialization in psychological research after World War II, it is a valid question to ask whether such a general theory of aesthetic experiences and underlying brain correlates should be the goal of the field. In my opinion, the striving for broad, overarching theories, even if they develop out of influential narrow theories such as those on aesthetic responses to human visual artworks, may yield crucial insights into the nature of aesthetic experiences that may otherwise have been missed. Even if extensive efforts devoted to advance a general theory of aesthetic experiences eventually fail, insights into the specificities of aesthetic experiences would not become apparent if aesthetic experiences were not systematically compared among different types of external objects (based on perception) and internal objects (based on thoughts and mental imagery). In other words, the question of whether specific aesthetic theories (e.g., those restricted to art) are justified can only be answered by trying to falsify the findings obtained for one type of object. As an illustration, inventive studies such as those comparing the processing of moral and facial beauty (Wang et al., 2015) or beauty induced by paintings and music (Ishizu and Zeki, 2011) should be highlighted. In a similar vein, neuroimaging studies investigating the beauty of mathematical formulas (Zeki et al., 2014) are equally encouraging.

How could such an ambitious goal of a general theory of neuroaesthetics be accomplished in the long run? One possibility is to further develop the study of other basic senses than vision, that is, to expand the *sensory-based* approach to empirical aesthetics. In general, it is not known yet whether effects reported for (different types of) visual objects generalize to other sensory domains. Therefore, a systematic comparative approach could offer new insights into meta-sensory findings and those that are unique for a sense. Such an approach does not always entail the use of different types of sensory stimuli within one research design, but also a replication of effects found for the visual domain with other sensory domains. For example, does the often reported mere-exposure effect (Zajonc, 1968) hold true for all types of sensory objects categories (Bornstein, 1989)? And more specifically, are the neural correlates of this effect, like those recently shown for faces (Kongthong et al., 2014), same or different across sensory domains and objects? Notably, the study of haptic perception in relation to aesthetic experiences has already yielded some promising results by following this research strand (Hintz and Nelson, 1971; Ballesteros and Reales, 2004; Jansson-Boyd and Marlow, 2007; Juricevic, 2009; Jakesch et al., 2011; Jakesch and Carbon, 2012; Etzi et al., 2014), which have recently not only led to initial speculations of underlying brain correlates (Gallace and Spence, 2011) but also to the development of a model of haptic aesthetic processing (Carbon and Jakesch, 2013).

As implied above, moving toward a general theory of neuroaesthetics will also require the systematic comparison of different object classes *within* one sensory domain, which is already ongoing to some degree (Vessel et al., 2014). However, there exists a vast array of possibilities of how to categorize objects within one sensory modality. In the visual domain, categories may range from human vs. non-human, human artwork vs. non-artwork, static vs. dynamic, utilitarian vs. decorative, to abstract vs. representational, to name only a few possible categories. The brain correlates of different object classes in relation to aesthetic experiences have already been partly compared in the visual domain (Cela-Conde et al., 2004; Kawabata and Zeki, 2004; Cattaneo et al., 2014), but this endeavor could be more rigorously followed and expanded to other sensory domains.

Another slowly growing research path is concerned with the direct comparison of perceptual, cognitive and affective processes *across* sensory domains within one research design. Such an approach is particularly valuable because it is mostly driven by concrete cross-domain research questions, often applies implicit or indirect tests, allows for a stricter control

of participant characteristics, and importantly, does not ignore the fact that we live in a multimodal world with multimodal artforms. In fact, there is a wide variety of recent behavioral and neurophysiological evidence against the notion of strict modularity of sensory modalities in humans (De Gelder and Bertelson, 2003; Shams and Kim, 2010; Gerdes et al., 2014) and animals (Van Wassenhove et al., 2012), which challenges the long-standing view of studying perception (and action) primarily from a unimodal perspective. In the field of empirical aesthetics, the small number of studies that follows this attitude does not take human artworks as a starting point for their investigations. Yet psychological concepts that have proven to be relevant for theories of aesthetic experiences, such as complexity (Boon et al., 2011; Marin and Leder, 2013), are currently examined across domains. Furthermore, aesthetic emotions in different artforms have been extensively studied (Silvia, 2005; Juslin, 2013; Koelsch, 2014a), but studies with crossmodal or multimodal designs are still outnumbered except for those on emotional pictures and sounds (e.g., Baumgartner et al., 2006a,b; Spreckelmeyer et al., 2006; Logeswaran and Bhattacharya, 2009; Marin and Bhattacharya, 2011; Petrini et al., 2011; Marin et al., 2012; Gerdes et al., 2014). A related issue concerns one of the core concepts of the field, namely beauty, and how it may differ in artforms (Augustin et al., 2012), types of objects (Marković, 2014), and neural processes (Ishizu and Zeki, 2011). Hopefully these valuable comparative studies will inspire researchers to continue working along these lines.

It must be remembered that a comparative approach to neuroaesthetics requires knowledge that may go beyond the expertise of an individual researcher, who usually works on a specialized research topic within a subfield of experimental psychology or cognitive neuroscience (e.g., emotion, visual, auditory and haptic perception, multisensory processing, psycho- and neurolinguistics, numerical cognition etc.). Clearly, collaborations of researchers with different backgrounds may yield better results. I also see much potential in a tighter collaboration between the humanities (musicology, art history, literary studies, dance studies, media and communication studies, semiotics, linguistics, philosophy, and mathematics) and the sciences (psychology, biology, and neuroscience) to gain a better understanding of aesthetic experiences. People trained in the humanities have the ability to detect aesthetically relevant phenomena by analytical thinking and abstraction. This usually goes hand in hand with the development of an appropriate terminology, classification system, and theoretical framework. For instance, the comparative study of human artforms has led to a research field called intermediality studies (Rajewsky, 2005), in which specific characteristics of media as well as their complex relations are the focus of interest. This field has already provided valuable theoretical insights into phenomena that contribute to people's aesthetic experiences across media such as music, literature and the visual arts (Wolf and Bernhart, 2006, 2007, 2013). However, these theories have not been subjected to empirical research methods yet. The humanities' awareness of the historical, social and cultural context in which artworks are embedded constitutes another asset that the sciences unfortunately often lack (Bullot and Reber, 2013; Redies, 2015). Importantly, intermediality studies could potentially be extended beyond the realm of arts and offer a refined alternative to the investigation of aesthetic experiences by going beyond the study of beauty and preference (Chatterjee and Vartanian, 2014; Consoli, 2015).

All things considered, I believe that advances in neuroaesthetics can be made by attracting more scientists and humanists from outside the visual aesthetics community, by comparing diverse aesthetic experiences across sensory modalities, and by giving up the notion that aesthetics concerns exclusively the study of beauty and the arts. Based on the knowledge generated by this comprehensive and comparative approach to neuroaesthetics, the study of top-down influences and moderators of aesthetic experiences as well as of aesthetic experiences induced by internal objects may become easier to accomplish from a neuroscientific perspective in the future.

Acknowledgments

This article was supported by the Open Access Publishing Fund of the University of Vienna. I would like to thank Bruno Gingras and Martina Jakesch for helpful comments on an earlier version of the manuscript.

References

Augustin, M. D., Carbon, C.-C., and Wagemans, J. (2012). Artful terms: a study on aesthetic word usage for visual art versus film and music. *Iperception* 3, 319–337. doi: 10.1068/i0511aap

Ballesteros, S., and Reales, J. M. (2004). Visual and haptic discrimination of symmetry in unfamiliar displays extended in the z-axis. *Perception* 33, 315–328. doi: 10.1068/p5017

Baumgarten, A. G. (2007, 1750/1758). *Aesthetica/Ästhetik, Vol. 2,* Latin text edited with facing German translation by ed D. Mirbach. Hamburg: Felix Meiner Verlag.

Baumgartner, T., Esslen, M., and Jäncke, L. (2006a). From emotion perception to emotion experience: emotions evoked by pictures and classical music. *Int. J. Psychophysiol.* 60, 34–43. doi: 10.1016/j.ijpsycho.2005.04.007

Baumgartner, T., Lutz, K., Schmidt, C. F., and Jäncke, L. (2006b). The emotional power of music: how music enhances the feeling of affective pictures. *Brain Res.* 1075, 151–164. doi: 10.1016/j.brainres.2005.12.065

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity.* New York, NY: McGraw-Hill.

Berlyne, D. E. (1971). *Aesthetics and Psychobiology.* New York, NY: Appleton.

Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., and Jacobs, A. M. (2013). When we like what we know - A parametric fMRI analysis of beauty and familiarity. *Brain Lang.* 124, 1–8. doi: 10.1016/j.bandl.2012.10.003

Boon, J. P., Casti, J., and Taylor, R. P. (2011). Artistic forms and complexity. *Nonlinear Dynamics Psychol. Life Sci.* 15, 265–283.

Bornstein, R. F. (1989). Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychol. Bull.* 106, 265. doi: 10.1037/0033-2909.106.2.265

Brattico, E., Bogert, B., and Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Front. Psychol.* 4:206. doi: 10.3389/fpsyg.2013.00206

Brattico, E., and Pearce, M. (2013). The neuroaesthetics of music. *Psychol. Aesthetics Creativity Arts* 7, 48–61. doi: 10.1037/a0031624

Brown, S., and Dissanayake, E. (2009). "The arts are more than aesthetics: neuroaesthetics as narrow aesthetics," in *Neuroaesthetics,* eds M. Skov and O. Vartanian (Amityville, NY: Baywood), 43–57.

Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., and Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *Neuroimage* 58, 250–258. doi: 10.1016/j.neuroimage.2011.06.012

Budd, M. (2002). *The Aesthetic Appreciation of Nature: Essays on the Aesthetics of Nature*. Oxford: Oxford University Press.

Bullot, N. J., and Reber, R. (2013). The artful mind meets art history: toward a psycho-historical framework for the science of art appreciation. *Behav. Brain Sci.* 36, 123–137. doi: 10.1017/S0140525X12000489

Calvo-Merino, B., Jola, C., Glaser, D. E., and Haggard, P. (2008). Towards a sensorimotor aesthetics of performing art. *Conscious. Cogn.* 17, 911–922. doi: 10.1016/j.concog.2007.11.003

Carbon, C.-C., and Jakesch, M. (2013). A model for haptic aesthetic processing and its implications for design. *Proc. IEEE* 101, 2123–2133. doi: 10.1109/JPROC.2012.2219831

Carlson, A. (2000). *Aesthetics and the Environment: The Appreciation of Nature, Art, and Architecture*. London: Routledge.

Cattaneo, Z., Lega, C., Gardelli, C., Merabet, L. B., Cela-Conde, C. J., and Nadal, M. (2014). The role of prefrontal and parietal cortices in esthetic appreciation of representational and abstract art: a TMS study. *Neuroimage* 99, 443–450. doi: 10.1016/j.neuroimage.2014.05.037

Cela-Conde, C. J., Agnati, L., Huston, J. P., Mora, F., and Nadal, M. (2011). The neural foundations of aesthetic appreciation. *Prog. Neurobiol.* 94, 39–48. doi: 10.1016/j.pneurobio.2011.03.003

Cela-Conde, C. J., Marty, G., Maestú, F., Ortiz, T., Munar, E., Fernández, A., et al. (2004). Activation of the prefrontal cortex in the human visual aesthetic perception. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6321–6325. doi: 10.1073/pnas.0401427101

Chatterjee, A., and Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn. Sci.*18, 370–375. doi: 10.1016/j.tics.2014.03.003

Chatterjee, A. (2004). Prospects for a cognitive neuroscience of visual aesthetics. *Bull. Psychol. Arts* 4, 56–60.

Chatterjee, A. (2011). Neuroaesthetics: a coming of age story. *J. Cogn. Neurosci.* 23, 53–62. doi: 10.1162/jocn.2010.21457

Chatterjee, A. (2013). *The Aesthetic Brain: How we Evolved to Desire Beauty and Enjoy Art*. New York, NY: Oxford University Press.

Chatterjee, A. (2014). "Neuroaesthetics: descriptive and experimental approaches," in *The Cambridge Handbook of the Psychology of Aesthetics and the Arts*, eds P. L. Tinio and J. K. Smith (Cambridge: Cambridge University Press), 481–499. doi: 10.1017/CBO9781139207058.023

Christensen, J. F., and Calvo-Merino, B. (2013). Dance as a subject for empirical aesthetics. *Psychol. Aesthetics Creativity Arts* 7, 76–88. doi: 10.1037/a0031827

Cinzia, D. D., and Vittorio, G. (2009). Neuroaesthetics: a review. *Curr. Opin. Neurobiol.* 19, 682–687. doi: 10.1016/j.conb.2009.09.001

Consoli, G. (2015). From beauty to knowledge: a new frame for the neuropsychological approach to aesthetics. *Front. Hum. Neurosci.* 9:290. doi: 10.3389/fnhum.2015.00290

Cross, E., and Ticini, L. (2012). Neuroaesthetics and beyond: new horizons in applying the science of the brain to the art of dance. *Phenomenol. Cogn. Sci.*11, 5–16. doi: 10.1007/s11097-010-9190-y

De Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014

Etzi, R., Spence, C., and Gallace, A. (2014). Textures that we like to touch: an experimental study of aesthetic preferences for tactile stimuli. *Conscious. Cogn.* 29, 178–188. doi: 10.1016/j.concog.2014.08.011

Gallace, A., and Spence, C. (2011). Tactile aesthetics: towards a definition of its characteristics and neural correlates. *Social Semiotics* 21, 569–589. doi: 10.1080/10350330.2011.591998

Gerdes, A. B. M., Wieser, M. J., and Alpers, G. W. (2014). Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains. *Front. Psychol.* 5:1351. doi: 10.3389/fpsyg.2014.01351

Grammer, K., Fink, B., Møller, A. P., and Thornhill, R. (2003). Darwinian aesthetics: sexual selection and the biology of beauty. *Biol. Rev.* 78, 385–407. doi: 10.1017/S1464793102006085

Hintz, J. M., and Nelson, T. M. (1971). Haptic aesthetic value of the golden section. *Br. J. Psychol.* 62, 217–223. doi: 10.1111/j.2044-8295.1971.tb02031.x

Ishizu, T., and Zeki, S. (2011). Toward a brain-based theory of beauty. *PLoS ONE* 6:e21852. doi: 10.1371/journal.pone.0021852

Jacobs, A. M. (2015). Neurocognitive Poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Front. Hum. Neurosci.* 9:186. doi: 10.3389/fnhum.2015.00186

Jacobsen, T. (2006). Bridging the arts and sciences: a framework for the psychology of aesthetics. *Leonardo* 39, 155–162. doi: 10.1162/leon.2006.39.2.155

Jakesch, M., and Carbon, C. C. (2012). The mere exposure effect in the domain of haptics. *PLoS ONE* 7:e31215. doi: 10.1371/journal.pone.0031215

Jakesch, M., Zachhuber, M., Leder, H., Spingler, M., and Carbon, C.-C. (2011). Scenario-based touching: on the influence of top-down processes on tactile and visual appreciation. *Res. Eng. Design* 22, 143–152. doi: 10.1007/s00163-010-0102-5

Jansson-Boyd, C., and Marlow, N. (2007). Not only in the eye of the beholder: tactile information can affect aesthetic evaluation. *Psychol. Aesthetics Creativity Arts* 1, 170–173. doi: 10.1037/1931-3896.1.3.170

Juricevic, I. (2009). Translating visual art into tactile art to produce equivalent aesthetic experiences. *Psychol. Aesthetics Creativity Arts* 3, 22–27. doi: 10.1037/a0014758

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Phys. Life Rev.* 10, 235–266. doi: 10.1016/j.plrev.2013.05.008

Kawabata, H., and Zeki, S. (2004). Neural correlates of beauty. *J Neurophysiol.* 91, 1699–1705. doi: 10.1152/jn.00696.2003

Koelsch, S. (2014a). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180. doi: 10.1038/nrn3666

Koelsch, S. (2014b). "From music perception to an integrative framework for the psychology of aesthetics," in *The Cambridge Handbook of the Psychology of Aesthetics and the Arts*, eds P. L. Tinio and J. K. Smith (Cambridge: Cambridge University Press), 300–336. doi: 10.1017/CBO9781139207058.016

Kongthong, N., Minami, T., and Nakauchi, S. (2014). Gamma oscillations distinguish mere exposure from other likability effects. *Neuropsychologia* 54, 129–138. doi: 10.1016/j.neuropsychologia.2013.12.021

Leder, H., and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode - Developments and challenges in empirical aesthetics. *Br. J. Psychol.* 105, 443–464. doi: 10.1111/bjop.12084

Logeswaran, N., and Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neurosci. Lett.* 455, 129–133. doi: 10.1016/j.neulet.2009.03.044

Marin, M. M., and Bhattacharya, J. (2011). "Music-induced emotions: some current issues and cross-modal comparisons," in *Music Education*, eds J. Hermida and M. Ferrero (Hauppauge, NY: Nova Science Publishers), 1–38.

Marin, M. M., Gingras, B., and Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion* 12, 618–631. doi: 10.1037/a0025020

Marin, M. M., and Leder, H. (2013). Examining complexity across domains: relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS ONE* 8:e72412. doi: 10.1371/journal.pone.0072412

Marković, S. (2014). Object domains and the experience of beauty. *Art Percept.* 2, 119–140. doi: 10.1163/22134913-00002020

Nadal, M., and Pearce, M. T. (2011). The copenhagen neuroaesthetics conference: prospects and pitfalls for an emerging field. *Brain Cogn.* 76, 172–183. doi: 10.1016/j.bandc.2011.01.009

Nadal, M., and Skov, M. (2013). Introduction to the special issue: toward an interdisciplinary neuroaesthetics. *Psychol. Aesthetics Creativity Arts* 7, 1. doi: 10.1037/a0031842

Nadal, M. (2013). "The experience of art: insights from neuroimaging," in *Progress in Brain Research, Vol. 204*, eds S. Finger, D. W. Zaidel, F. Boller, and J. Bogousslavsky (Amsterdam: Netherlands), 135–158. doi: 10.1016/B978-0-444-63287-6.00007-5

Petrini, K., Crabbe, F., Sheridan, C., and Pollick, F. E. (2011). The music of your emotions: neural substrates involved in detection of emotional correspondence between auditory and visual music actions. *PLoS ONE* 6:e19165. doi: 10.1371/journal.pone.0019165

Rajewsky, I. O. (2005). Intermediality, intertextuality, and remediation: a literary perspective on intermediality. *Intermédialités* 6, 43–64. doi: 10.7202/1005505ar

Ramachandran, V. S., and Hirstein, W. (1999). The science of art: a neurological theory of aesthetic experience. *J. Conscious. Stud.* 6, 15–51.

Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Front. Hum. Neurosci.* 9:218. doi: 10.3389/fnhum.2015.00218

Shams, L., and Kim, R. (2010). Crossmodal influences on visual perception. *Phys. Life Rev.* 7, 269–284. doi: 10.1016/j.plrev.2010.04.006

Shimamura, A. P., and Palmer, S. E. (2012). *Aesthetic Science: Connecting Minds, Brains, and Experience.* New York, NY: Oxford University Press.

Silvia, P. J. (2005). Emotional responses to art: from collation and arousal to cognition and emotion. *Rev. Gen. Psychol.* 9:342. doi: 10.1037/1089-2680.9.4.342

Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., and Münte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Res.* 1070, 160–170. doi: 10.1016/j.brainres.2005.11.075

Starr, G. G. (2013). *Feeling Beauty: The Neuroscience of Aesthetic Experience.* Cambridge: MIT Press.

Tinio, P. P. L. (2013). From artistic creation to aesthetic reception: the mirror model of art. *Psychol. Aesthetics Creativity Arts* 7, 265–275. doi: 10.1037/a0030872

Van Wassenhove, V., Ghazanfar, A., Munhall, K., Schroeder, C., and Hall, G. (2012). "Bridging the gap between human and non human studies of audiovisual integration," in *The New Handbook of Multisensory Processing,* ed B. E. Stein (Cambridge, MA: MIT Press), 153–167.

Vessel, E. A., Stahl, J., Maurer, N., Denker, A., and Starr, G. G. (2014). "Personalized visual aesthetics," in *Proceedings. SPIE 9014, Human Vision and Electronic Imaging XIX, 90140S*, eds B. E. Rogowitz, T. N. Pappas, and H. de Ridder (San Francisco, CA), 1–8. doi: 10.1117/12.2043126

Wang, T., Mo, L., Mo, C., Tan, L. H., Cant, J. S., Zhong, L., et al. (2015). Is moral beauty different from facial beauty? Evidence from an fMRI study. *Soc. Cogn. Affect. Neurosci.* 10, 814–823. doi: 10.1093/scan/nsu123

Wolf, W., and Bernhart, W. (eds.). (2006). *Framing Borders in Literature and Other Media.* Amsterdam; New York: Rodopi.

Wolf, W., and Bernhart, W. (eds.). (2007). *Description in Literature and Other Media.* Amsterdam; New York: Rodopi.

Wolf, W., and Bernhart, W. (eds.). (2013). *Immersion and Distance: Aesthetic Illusion in Literature and Other Media.* Amsterdam; New York: Rodopi.

Zaidel, D. W. (2015). Neuroaesthetics is not just about art. *Front. Hum. Neurosci.* 9:80. doi: 10.3389/fnhum.2015.00080

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9, 1. doi: 10.1037/h0025848

Zatorre, R. J., and Salimpoor, V. N. (2013). From perception to pleasure: music and its neural substrates. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10430–10437. doi: 10.1073/pnas.1301228110

Zeki, S., and Nash, J. (1999). *Inner Vision: An Exploration of Art and the Brain.* Oxford: Oxford University Press.

Zeki, S., Romaya, J. P., Benincasa, D. M. T., and Atiyah, M. F. (2014). The experience of mathematical beauty and its neural correlates. *Front. Hum. Neurosci.* 8:68. doi: 10.3389/fnhum.2014.00068

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Marin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

2.3 Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings (Marin & Leder, 2016)

2.3.1 Introduction

Aesthetic experiences are thought to emerge over time (e.g., Belfi et al., 2019; Brattico et al., 2013; Carbon, 2012; Graf & Landwehr, 2015; Leder et al., 2004; Muth et al., 2015; Pelowski et al., 2020), but little is known about the dynamics of aesthetic experiences (Strijbosch et al., 2022). Similarly, the time course of complexity perception has not been studied in great depth (Bachmann & Vipper, 1983; Cupchik & Gebotys, 1988), let alone how the association between objective and subjective complexity may change over time. Thus, the main goal of the current study, which consists of three experiments and follows a comparative approach as proposed in Marin and Leder (2013) and Marin (2015), was to examine whether and how subjective ratings of familiarity, complexity, arousal and pleasantness differ across three presentation durations for environmental scenes and representational paintings. We were also interested in how the interrelationships between these variables change over time, and further, how presentation duration affects the relationship between subjective and objective measures of complexity (Experiments 1 and 2). Regarding Berlyne's theory of collative motivation (1971), we aimed at testing mediation effects of arousal across all three presentation durations. For this purpose, we used the very same stimulus material of Marin and Leder (2013) and employed the same objective measures. We chose presentation durations of 1, 5 and 25 s with the goal of covering frequently used presentation durations in visual perception science (i.e., short, moderate and long). A presentation duration of 25 s is close to the average time spent in front of paintings in museums (Brieber et al., 2014; Smith & Smith, 2001). Our experiments only involved female participants due to gender effects reported in Marin and Leder (2013) and the smaller number of available male psychology students.

Another goal was to study the time course of the multidimensionality of complexity (Nadal et al., 2010). Whereas general complexity ratings were collected in Experiments 1 and 2, Experiment 3 asked for three specific ratings that may underlie complexity perception based on the surface structure of a visual scene, namely ratings of figure-ground organization vs. complex scenes (a measure that has been previously used for environmental scenes, i.e., IAPS pictures), the number of elements and the degree of disorganization. We were thus able to examine how these dimensions of visual complexity are perceived when presented for 1 and 5 s, and also, how

these dimensions were correlated with objective measures of complexity. Moreover, it was possible to determine how these dimensions of complexity were related to general complexity ratings collected in Experiments 1 and 2. Again, we compared our results across stimulus categories to increase our understanding of the characteristics of visual aesthetic experiences.

2.3.2 Authorship contribution statement

MMM Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft preparation, visualization, and project administration. HL Supervision, resources, writing − reviewing and editing.

2.3.3 Re-print of the published research article

Acta Psychologica 163 (2016) 38–58

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings

Manuela M. Marin *, Helmut Leder

Department of Basic Psychological Research and Research Methods, University of Vienna, Austria

ARTICLE INFO

Article history:
Received 18 April 2015
Received in revised form 15 July 2015
Accepted 11 October 2015
Available online xxxx

Keywords:
Empirical esthetics
Exposure time
Complexity
IAPS
Visual arts
Objective measures

ABSTRACT

Complexity constitutes an integral part of humans' environment and is inherent to information processing. However, little is known about the dynamics of visual complexity perception of affective environmental scenes (IAPS pictures) and artworks, such as affective representational paintings. In three experiments, we studied the time course of visual complexity perception by varying presentation duration and comparing subjective ratings with objective measures of complexity. In Experiment 1, 60 females rated 96 IAPS pictures, presented either for 1, 5, or 25 s, for familiarity, complexity, pleasantness and arousal. In Experiment 2, another 60 females rated 96 representational paintings. Mean ratings of complexity and pleasantness changed according to presentation duration in a similar vein in both experiments, suggesting an inverted U-shape. No common pattern of results was observed for arousal and familiarity ratings across the two picture sets. The correlations between subjective and objective measures of complexity increased with longer exposure durations for IAPS pictures, but results were more ambiguous for paintings. Experiment 3 explored the time course of the multidimensionality of visual complexity perception. Another 109 females rated the number of objects, their disorganization and the differentiation between a figure-ground vs. complex scene composition of pictures presented for 1 and 5 s. The multidimensionality of visual complexity only clearly emerged in the 5-s condition. In both picture sets, the strength of the correlations with objective measures depended on the type of subdimension of complexity and was less affected by presentation duration than correlations with general complexity in Experiments 1 and 2. These results have clear implications for perceptual and cognitive theories, especially for those of esthetic experiences, in which the dynamical changes of complexity perception need to be integrated.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Visual complexity perception and changes over time

Measuring the time course of perceptual and cognitive processes is generally regarded as a central issue of experimental psychology. Not surprisingly so, visual art perception has been modeled as a dynamic process that involves multiple perceptual, cognitive and affective processing stages that interact over time (Carbon, 2012; Chatterjee, 2003; Leder, Belke, Oeberst, & Augustin, 2004; Tinio, 2013). However, the field of empirical esthetics has, for the most part, neglected the detailed study of the time course of these processes leading to esthetic judgments, visual evaluations such as complexity, and emotions (Chatterjee, 2014; Jacobsen, 2014; Leder & Nadal, 2014). Granted that the methodological difficulties in such an endeavor are manifold, the lack of knowledge of the dynamics of esthetic processes is rather unexpected given that, for example, the time course of emotional processing is commonly studied on a behavioral and neural level (e.g., Borgomaneri, Gazzola, & Avenanti, 2014; Codispoti, Mazzetti, & Bradley, 2009; Golland, Keissar, & Levit-Binnun, 2014; Martini et al., 2012). Moreover, pattern complexity and its perceptual time course have been studied within psychophysics (e.g., Chipman, 1977; Ichikawa, 1985; Terwilliger, 1963; Vitz, 1966), but research involving more ecologically-valid stimuli is scarce. In this study, we thus investigated the perceptual dynamics of visual complexity perception in affective environmental scenes and representational paintings by varying the presentation duration and combining subjective ratings with objective measures of complexity.

Complexity is widely studied in psychology and related disciplines due to its central role in humans' environment (e.g., Kaplan & Kaplan, 1982; Schooler, 1984). Therefore, the behavioral outcomes of sensory, cognitive and affective mechanisms elicited by visual and auditory stimuli varying in perceived complexity have been a common research topic over the last decades (e.g., Aksentijevic & Gibson, 2012; Bradley, Houbova, Miccoli, Costa, & Lang, 2011; Donderi, 2006a; Gibson, 1998; Heaps & Handel, 1999; Eerola, Himberg, Toiviainen, & Louhivuori, 2006). For example, research in empirical esthetics has been largely motivated by the question of what determines humans' preferences for certain kinds of man-made or natural stimuli over others. One

* Corresponding author at: Liebiggasse 5, A-1010 Vienna, Austria.
E-mail address: manuela.marin@univie.ac.at (M.M. Marin).

http://dx.doi.org/10.1016/j.actpsy.2015.10.002
0001-6918/© 2015 Elsevier B.V. All rights reserved.

influential model was developed by Daniel Berlyne (Berlyne 1966; Berlyne, 1960, 1971, 1974), whose collative-motivation model, based on Wundt's curve (1874), posits that people generally prefer stimuli of intermediate complexity to simple and highly complex ones under normal arousal conditions. Intermediate levels of complexity give rise to an optimal level of arousal, yielding the highest hedonic value. According to Berlyne, an increase in complexity leads to the activation of a primary reward system up to an optimum. An aversion system becomes active upon a further increase in complexity, yielding an inverted U-shaped function between complexity and hedonic value. Collative variables are related to the viewer's expectations and comprise stimulus factors such as complexity, novelty, uncertainty, conflict and unfamiliarity. Although Berlyne's model has not been consistently empirically verified and may be debatable on theoretical grounds (Hekkert & Wieringen, 1990; Marin & Leder, 2013; Martindale, Moore, & Borkum, 1990), complexity has been reliably identified as a determinant of visual preferences and has therefore been incorporated into current models of esthetic experiences (Chatterjee, 2003; Leder et al., 2004; Locher, Overbeeke, & Wensveen, 2010; Tinio, 2013). However, if the evaluation of basic visual properties such as complexity systematically evolves over time, models relating complexity to esthetic appeal need to consider this change.

Most relevant to the study at hand are the few studies on visual complexity employing a set of different presentation durations within the context of Berlyne's (1971) psychobiological model (Cupchik & Berlyne, 1979; Cupchik & Gebotys, 1988), and those studies using semantic differential ratings of paintings (Bachmann & Vipper, 1983; Smith, Bousquet, Chang, & Smith, 2006). Specifically, Cupchik and Berlyne (1979) asked how quickly collative properties of stimuli (complexity, uncertainty and redundancy) can be discriminated by the observer, and further, whether the perception of these structural properties changes as a function of presentation duration. In two experiments, the authors followed a synthetic (i.e., manipulation of simple artificial materials) and analytic (i.e. use of genuine works of art) approach to the study of esthetic experiences (Berlyne, 1971). In Experiment 1, the number of elements and their organization were manipulated in a black and white checkerboard pattern. Moreover, twelve paintings were chosen to vary in complexity and/or arousal. Stimuli were presented for 50 ms, 500 ms and 5000 ms, with presentation duration as a within-subject factor, and participants completed five ratings referring to complexity, orderliness (organization), pleasingness (hedonic tone) and arousal. The results indicated that collative properties could be discriminated after a single glance in genuine artistic and artificial materials. The results also showed a monotonic increase of orderliness as a function of presentation duration. Pleasingness ratings were lower for the 50-ms condition than for longer presentation durations. For both paintings and patterns, mean ratings of scales referring to arousal were higher for the 50-ms condition than for the two longer conditions, which did not differ in their mean ratings. The authors concluded that the increased perception of order for the longer presentation durations was related to higher levels of pleasingness.

In a similar vein, Bachmann and Vipper (1983) were interested in the microgenesis of perceptual ratings of paintings from different artistic styles. Six different schools of arts, ranging from realism to surrealism, and four different presentation durations were used: 1 ms, 20 ms, 100 ms and 500 ms. The factor of presentation duration affected ratings scales such as “involved-indifferent”, “simple-complex”, “regular-chaotic” and “precise-vague” in a way that suggests that the perception of paintings develops from a general to a more differentiated character. For instance, paintings were rated as more complex when presentation durations were short compared to longer exposures. This pattern of results was evident for all four scales and was only slightly affected by artistic style.

Although the examination of early perceptual dynamical processes is particularly fruitful within the context of Gestalt formation (Bachmann & Vipper, 1983), the study of longer presentation durations may have greater relevance in real life situations. Smith and Smith (2001) found the amount of time spent in front of artworks to be on average 27.2 s (i.e., looking at the painting and its label), with a median viewing time of 17 s. In a rare comparison across a wide range of presentation times, the study by Smith et al. (2006) presented artworks for 1, 5, 30, or 60 s. Four paintings (2 impressionists, 2 modern) were shown in one of the four time conditions and either presented with or without labels. Participants with limited art experience rated the artworks according to 24 adjective pairs, yielding three subscales: evaluative, activity and potency/structure. The authors ran separate analyses for each painting and did not find any significant effects of presentation durations on subjective ratings. However, due to the small set of selected paintings it is possible that the results were largely dependent on the specific paintings and thus cannot be generalized.

Visual complexity is not a uniform concept. The multidimensional nature of visual complexity has been acknowledged with regard to natural scenes, abstract patterns, and artworks and was described by factors such as the number and variety of elements, their disorganization, as well as the degree of symmetry (e.g., Berlyne, Ogilvie, & Parham, 1968; Chipman, 1977; Ichikawa, 1985; Nadal, Munar, Marty, & Cela-Conde, 2010). The factor relating to the number and variety of elements seems to be the strongest predictor of visual complexity (Berlyne et al., 1968; Nadal et al., 2010) and may have a different perceptual time course than structural factors (Ichikawa, 1985). Symmetry is sometimes studied as an independent factor of complexity when abstract patterns, simple figures or faces are investigated (e.g., Day, 1967; Eisenman, 1968; Gartus & Leder, 2013; Jacobsen & Höfel, 2002; Tinio, Gerger, & Leder, 2013). However, across the above-mentioned studies the presentation durations varied considerably, making it difficult to compare across stimulus sets. Thus, it is unclear whether the relationship between general visual complexity and its subdimensions changes over time.

1.2. Comparing complexity perception of artistic and non-artistic stimuli

The comparison of esthetic responses to different image categories has attracted wide interest in empirical esthetics (Marković, 2014; Vessel & Rubin, 2010; Vessel, Stahl, Maurer, Denker, & Starr, 2014) and in computational esthetics (Braun, Amirshahi, Denzler, & Redies, 2013; Graham & Field, 2007; Graham & Meng, 2011; Redies, Amirshahi, Koch, & Denzler, 2012). This is probably due to the fact that such a comparison may yield insights into questions that lie at the heart of visual esthetics: What is special about the perception of human-made artworks in comparison to natural scenes, landscapes and faces? And how and why do humans'esthetic responses to diverse visual stimuli differ from each other? These questions may for example be addressed by a systematic comparison of the relationships of variables defining Berlyne's psychobiological model of esthetic experience (1971) in pictures that are either photographs of human-made artworks, i.e. products of human creativity and craftsmanship meant to induce esthetic experiences, or photographs of environmental scenes evoking affective and motivational responses in humans. Marin and Leder (2013) compared subjective esthetic responses (familiarity, pleasantness, arousal and complexity) and objective measures of complexity in two sets of affective visual stimuli (representational artworks and environmental scenes presented for 25 s), both comprising a similar range of semantic contents. They found that the emotion space, spanned by pleasantness and arousal, was similar in environmental scenes and representational paintings. Furthermore, they reported a stronger relationship between complexity and arousal in artworks than in environmental scene photographs. In both types of visual stimuli, the relationship between pleasantness and complexity did not show an inverted U-shape as predicted by Berlyne (1971). Instead, the two variables were negatively related. Arousal was identified as a mediator in the complexity–pleasantness relationship in environmental scenes and representational paintings. Marin and Leder (2013) also found

several gender effects in their analyses, highlighting the need to consider gender in the study of esthetic responses to visual scenes with emotional content.

Affective response patterns to environmental scenes, such as to those of the International Affective Picture System (IAPS) (Lang, Bradley, & Cuthbert, 2005), are usually studied when pictures are shown for a sustained time period of 6 s because normative ratings were obtained for this presentation duration. Studies probing affective and motivational reactions to IAPS pictures presented for shorter or longer durations than 6 s are rare. Codispoti et al. (2009) examined subjective responses to neutral, pleasant and unpleasant IAPS pictures which were presented by using a wide range of presentation durations, ranging from 25 to 6000 ms. Their findings indicated that pleasantness ratings of unpleasant pictures in the 25- and 80-ms conditions were higher than in the longer conditions. Similarly, pleasant pictures were rated as less pleasant in the 25-ms condition than in the longer time conditions. Arousal ratings of pleasant and unpleasant pictures were also lower in the 25-ms condition than in longer presentation durations. The graphical representation of the pleasantness-arousal space across different presentation durations showed that the subjective ratings of pictures become more differentiated the longer the presentation duration. In other words, the emotion space expands over time. Taken together, these findings constitute tentative evidence that changes of feeling states with regard to environmental scenes may occur at very early processing stages but become more stable after longer exposures to a stimulus.

1.3. Objective measures of visual complexity

Recent advances in the study of subjective complexity and its relationship to esthetic experiences have included the use of objective measures of complexity (Forsythe, 2009; Forsythe, Nadal, Sheehy, Cela-Conde, & Sawey, 2011; Marin & Leder, 2013; Tuch, Bargas-Avila, Opwis, & Wilhelm, 2009). These measures were based on various image compression techniques (Bates et al., 2003; Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012; Donderi, 2006a; Forsythe, Mulhern, & Sawey, 2008; Forsythe, Sheehy, & Sawey, 2003; Palumbo, Odgen, Makin, & Bertamini, 2014) or edge detection algorithms (Cavalcante et al., 2014; Chikhman et al., 2012; Forsythe et al., 2008; Forsythe et al., 2011) and mostly yielded moderate correlations with subjective complexity ratings (i.e., general complexity). For example, Forsythe et al. (2011) computed compressed file size and perimeter detection of five sets of visual stimuli shown for 5 s. The objective measures correlated moderately to highly with subjective ratings of complexity, the degree of correlation varying across the stimulus sets (abstract artistic, abstract decorative, figurative, figurative decorative and natural pictures). Marin and Leder (2013) extended this line of research by applying a larger set of objective complexity measures to affective environmental scenes and representational artworks. They found that the performance of an objective complexity measure depended on the specific stimulus set, which is in line with a study involving Chinese hieroglyphs and outlines of common objects (Chikhman et al., 2012). In these and other studies employing objective measures of complexity (e.g., Cavalcante et al., 2014; Forsythe et al., 2008; Donderi, 2006a; Palumbo et al., 2014) the stimuli were presented for several seconds and much longer, unrestricted exposure times. However, it is yet unclear whether the performance of an objective complexity measure depends on the presentation duration of the stimulus.

1.4. The present study

Here, we sought to further probe the time course of visual complexity processing by following a comparative approach within the context of the study by Marin and Leder (2013). In Experiments 1 and 2, large sets of IAPS pictures (i.e. environmental scenes) and affective representational artworks (taken from Marin and Leder (2013)) were either presented for 1, 5 or 25 s and rated for familiarity, complexity, pleasantness and arousal by six groups of participants. This allowed us to ascertain whether the dynamics of complexity perception regarding affective stimuli can be generalized across stimulus types. Furthermore, we included longer presentation durations because they capture real-life esthetic episodes to a greater degree than the study of the microgenesis of early perceptual processes. Changes in subjective experiences of complexity after longer presentation durations were difficult to predict. However, based on findings involving sets of short presentation durations (50 ms–5 s) (Bachmann & Vipper, 1983; Cupchik & Berlyne, 1979), it was surmised that subjective complexity ratings of the visual input would decrease with presentation duration. An alternative hypothesis stated that longer presentation durations would yield higher perceived complexity because more details of the images may become apparent and top-down cognitive influences regarding the semantic content may occur and interact with the perceived complexity of the structure of the visual signal. Furthermore, we decided to confine our sample to females since various types of gender effects with regard to emotional processing of IAPS pictures and representational artworks have previously been reported, as shown by subjective ratings and neurophysiological responses (e.g., Bradley, Codispoti, Sabatinelli, & Lang, 2001; Lithari et al., 2010; Marin & Leder, 2013).

Besides the study of complexity perception including longer presentation durations, another noteworthy aspect of Experiments 1 and 2 was the inclusion of objective complexity measures in the study of the effect of presentation duration on complexity judgments, following the tradition of psychophysics. If objective measures of complexity are able to capture the complexity with regard to the visual structure of an image to some extent, their correlation strengths with subjective ratings across different presentation durations are a valuable means to study bottom-up versus top-down contributions to subjective complexity. For example, an increase in correlation strength across the three presentation durations would indicate that subjective ratings are related more strongly to the visual structure of the image at longer presentation durations than at shorter presentation durations. In other words, the visual system is better able to judge the objective complexity of an image. A decrease in correlation strength would suggest the reverse and indicate that higher-order processes interact to a stronger degree with subjective ratings at longer than at shorter presentation durations. Another possibility was that the results will not show a pattern of linear increase or decrease but a quadratic relationship. We used a large set of objective measures (see Marin and Leder (2013), four compressed file size measures (JPEG, GIF, PNG and TIFF), Canny and perimeter edge detection, entropy, three RMS contrast and three phase congruency measures) in order to be able to see whether general trends within and across stimulus sets, independent of the performance of a specific measure, could be detected. Experiments 1 and 2 also offered novel insights into whether the mediating effect of arousal in Berlyne's psychobiological model (1971) depended on presentation duration, and whether the relationships between complexity, arousal and pleasantness were similar for environmental scenes and representational paintings.

In Experiment 3, we investigated the time course of the multidimensionality of visual complexity in comparison to general complexity ratings obtained in Experiments 1 and 2. For this purpose, three scales were used to rate environmental scenes and representational paintings presented for 1 and 5 s by two independent groups: figure-ground organization vs. complex scene, the number of elements, and their degree of (dis)organization (see Nadal et al. (2010)). We predicted a decrease of disorganization ratings (Cupchik & Berlyne, 1979). Moreover, as the relationships between objective complexity measures and subjective ratings of subdimensions of complexity have not been investigated so far, this experiment allowed us to investigate which subjective ratings of the subdimensions of complexity yield stronger, or weaker, correlations with objective measures than ratings of general complexity. Finally, we sought to find out how these correlations depend on the stimulus type and change according to the presentation durations.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixty German-speaking female psychology undergraduates participated in exchange for course credit. All participants had normal or corrected-to-normal visual acuity. Twenty participants were tested in one of the three conditions of presentation duration (1 s, 5 s and 25 s). The mean age of these three participant groups (1 s: 21.9 ± 2.3 years; 5 s: 22.2 ± 3.0 years; 25 s: 23.5 ± 4.4 years) did not differ significantly, $H(2) = 2.26$, $p = .323$. For the 25-s condition, data of 18 participants were taken from Experiment 1 as reported in Marin and Leder (2013).

2.1.2. Materials

The set of 96 pictures selected from the IAPS (Lang et al., 2005) was identical to the one used by Marin and Leder (2013). Colored pictures of natural scenes were chosen on the basis of their standardized values of arousal and pleasantness to vary in hedonic content (low- vs. high-arousing and unpleasant vs. pleasant). Following Russell's circumplex model of affect (1980), this pre-selection led to 24 pictures of specific hedonic contents (either low-arousing pleasant, high-arousing pleasant, low-arousing unpleasant, or high-arousing unpleasant). The results reported by Marin and Leder (2013) for a presentation duration of 25 s indicated that the 96 pictures followed a U-shaped distribution in the 2-dimensional pleasantness-arousal space, which suggests that the pre-selected pictures largely covered the expected affective space induced by this type of environmental scene photographs (Lang et al., 2005).

Half of the pictures were chosen to depict rather simple figure-ground compositions (i.e., one figure with a uniform background), and the other half depicted complex scenes with several objects and a more varied background (see Bradley et al. (2007, 2011)). The semantic contents comprised animals, plants, landscapes, food and human beings in everyday life scenes, but no pictures showing erotic scenes, brand names and strong mutilation. All pictures were in landscape format (1024 × 768 pixels) and saved in their original JPEG (Joint Photographic Expert Group) format.

2.1.3. Procedure

The experimental procedure of all three Experiments followed the current version of the Ethical Principles of Psychologists and Code of Conduct of the American Psychological Association (with 2010 amendments). Participants read and signed an informed consent form prior to participating. The procedure of the experiment was identical to the one described in Experiment 1 in Marin and Leder (2013), with the exception that participants were randomly assigned to either the 1- or 5-s condition of presentation duration. Participants were familiarized with the task in two practice trials. Each picture was announced on the screen (Samsung SyncMaster S2443BW, 24 in.) and followed after 5 s on a black background. Participants, sitting 60–70 cm away from the screen, were instructed to look at the picture for the entire presentation duration. After the pictures had disappeared, they were rated on a seven-point Likert scale in terms of familiarity ("Please rate your familiarity with the contents of the picture", with "unfamiliar" and "very familiar" as anchors), complexity ("Please rate your felt complexity of the picture", with "very simple" and "very complex" as anchors), felt pleasantness ("Please rate the degree of pleasantness of your emotional experience", with "very unpleasant" and "very pleasant" as anchors) and felt arousal ("Please rate your felt arousal", with "very calm" and "very excited" as anchors). The order of the ratings was kept constant across participants. Familiarity was rated first because it is known that familiarity is related to most affective and esthetic measures. Complexity was rated second because it was of main interest to the researchers. It was ensured that participants understood that complexity refers to the visual structure of the image. Pleasantness and arousal, as two primary dimensions of emotional response, were rated last because these measures were of secondary interest. Note that unpublished data based on females who rated the same stimuli for a presentation duration of 5 s with a different rating order (pleasantness–arousal–complexity–familiarity) yielded similar effects regarding the interrelationships of these variables, suggesting that order effects are unlikely. Participants indicated their ratings by mouse click after which the next rating scale appeared on the screen. Importantly, participants were instructed to report their subjective feelings rather than the perceived emotional contents of the pictures (felt vs. perceived emotion, see Gabrielsson (2002)) and to give all ratings spontaneously (although they were informed that there were no time constraints).

The pictures were presented in four blocks and the order of the blocks was randomized across participants. Within each block, 24 pictures of similar hedonic content (either low-arousing pleasant, high-arousing pleasant, low-arousing unpleasant, or high-arousing unpleasant) but with varying complexity were randomly presented. This approach encouraged participants to perceive subtle differences in terms of complexity and emotional contents. Participants were allowed to take a short and self-paced break between these blocks and told that the four blocks would vary in terms of the images' emotional contents. The length of the experiment depended on the presentation duration and varied between 30 and 60 min. After the experiment, participants were thanked, debriefed and dismissed.

The very same set of 13 objective complexity measures as those reported in Marin and Leder (2013) were used to analyze the photographs, including file sizes of four compression formats (JPEG, GIF, PNG and TIFF), several measures of edge detection (perimeter detection, Canny edge detection (Canny, 1986), root mean square (RMS) contrast (Cavalcante et al., 2014; Peli, 1990) and edge detection based on phase congruency (Kovesi, 1999, 2003) as well as the entropy of a gray-scale image (Gonzalez, Woods, & Eddins, 2003). Detailed information about how these measures were computed is available in Marin and Leder (2013).

2.2. Results

The results section comprises three parts. First, findings regarding the change of subjective experience over time are presented, including aspects such as the shape and size of the 2-dimensional emotion space, the inter-relationships between the subjective ratings as well as the change of averaged ratings across all pictures over time. Second, results of mediated regressions are presented to test whether Berlyne's collative motivation model may depend on presentation duration. Third, for all presentation durations the relationships between objective measures of complexity and subjective ratings, notably with subjective complexity, are shown.

2.2.1. Change of subjective ratings over time

Prior to any further analyses, subjective ratings for each picture in each condition of presentation duration were averaged across participants. Inter-rater reliability was assessed by calculating the average intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency (McGraw & Wong, 1996; Shrout & Fleiss, 1979). We observed a very high inter-rater reliability for all four types of ratings in all three conditions of presentation duration. For the presentation duration of 1 s, the results were as follows: familiarity ($ICC(2, 20) = .93$, 95% confidence interval (CI) [.91, .95]), complexity ($ICC(2, 20) = .84$, 95% CI [.79, .88]), pleasantness ($ICC(2, 20) = .97$, 95% CI [.96, .98]), and arousal ($ICC(2, 20) = .88$, 95% CI [.84, .91]). Similar results were obtained for the presentation duration of 5 s: familiarity ($ICC(2, 20) = .93$, 95% CI [.91, .95]), complexity ($ICC(2, 20) = .86$, 95% CI [.82, .90]), pleasantness ($ICC(2, 20) = .98$, 95% CI [.97, .98]), and arousal ($ICC(2, 20) = .86$, 95% CI [.82, .90]). Inter-rater reliability was also considered as sufficiently high for the 25-s condition: familiarity ($ICC(2, 20) = .93$, 95% CI [.91, .95]), complexity ($ICC(2, 20) = .89$, 95% CI [.86, .92]), pleasantness ($ICC(2, 20) = .98$, 95% CI [.97, .98]), and

arousal (ICC(2, 20) = .89, 95% CI [.85, .92]). Since the inter-rater reliability of each variable was very similar across the three presentation durations, the comparison of the pictures' ratings across the three conditions was justified.

An exploratory data analysis revealed that each of the variables referring to the different compression formats and other measures of objective complexity (edge detection, RMS contrast, entropy) contained several outliers that were 2 *SD* above or below the mean. A series of Kolmogorov–Smirnov normality tests indicated that the data, after removing the outliers, of the different compression formats and measures of edge detection, RMS contrast and entropy deviated significantly from a normal distribution. Furthermore, subjective ratings of familiarity, complexity, pleasantness and arousal were generally not normally distributed, and these distributions did not benefit from transformations. Thus, it was decided to remove outliers and to employ non-parametric analyses to test for differences between conditions and non-parametric correlational analyses to investigate the inter-relationships between the sets of variables[1,2]

In Fig. 1a–c, the distribution of IAPS pictures in the arousal–pleasantness emotion space is depicted for three presentation durations. In general, the three distributions do not markedly differ from each other. The current subset of the IAPS picture system followed the previously reported quadratic relationship between pleasantness and arousal (e.g., Codispoti et al., 2009; Lang et al., 2005; Marin & Leder, 2013) in all three presentation duration conditions. Curve fitting analyses were conducted to confirm the quadratic relationship between arousal and pleasantness. The quadratic relationship in the 1-s condition, $R^2 = .69$, $F(2, 92) = 104.01$, $p < .001$, $y = -2.47x + .28x^2 + 8.84$, was somewhat stronger than in the 5-s condition, $R^2 = .62$, $F(2, 92) = 75.45$, $p < .001$, $y = -2.41x + .29x^2 + 8.69$. The relationship for the 25-s condition was similar to the one of the 5-s condition, $R^2 = .62$, $F(2, 93) = 75.27$, $p < .001$, $y = -2.80x + .34x^2 + 9.29$. Furthermore, Spearman's rank order correlations were conducted between the three arousal ratings (one per presentation duration) with the following results: $r_s(92) = .80$ for $arousal_{1s}$ – $arousal_{5s}$, $r_s(90) = .84$ for $arousal_{1s}$ – $arousal_{25s}$, and $r_s(89) = .89$ for $arousal_{5s}$ – $arousal_{25s}$. Somewhat higher correlations were observed for pleasantness ratings: $r_s(94) = .97$ for $pleasantness_{1s}$ – $pleasantness_{5s}$, $r_s(94) = .96$ for $pleasantness_{1s}$ – $pleasantness_{25s}$, and $r_s(94) = .98$ for $pleasantness_{5s}$ – $pleasantness_{25s}$. These results suggest that pleasantness ratings may be less affected by presentation duration than arousal ratings.

In order to allow for a more comprehensive, global comparison between the different emotion spaces, three Procrustes analyses (Bookstein, 1997; Kendall, 1984) based on the averaged pleasantness and arousal ratings were performed in Matlab (The MathWorks, Inc., Natick, Massachusetts, USA). Procrustes analysis enables the comparisons of two sets of data by least-squares orthogonal mapping and provides a descriptive summary. Two common measures reported in the examination of the similarity of the two shapes are a standardized dissimilarity measure *d* (ranging between 0 and 1 with lower numbers indicating higher similarity) and a scaling measure *b* that refers to the size similarity of the two shapes. The comparison between the 1 s- and 5-s conditions revealed that the two emotions spaces were similar in shape ($d = .118$) and size ($b = .865$), which was also the case for the comparison between the 1-s and 25-s conditions ($d = .112$, $b = .825$). However, the smaller *b* value in the latter comparison suggests that the size of the emotion spaces of the 1-s and 25-s conditions is more dissimilar from each other than the size of the emotion spaces of the 1-s and 5-s conditions. The final Procrustes analysis compared the emotion spaces of the 5-s and the 25-s conditions and the results indicated the highest similarity in shape ($d = .085$) and size ($b = .909$) among the three comparisons. This finding is corroborated by the visual representation of the data (Fig. 1), the curve-fitting analyses as well as by the comparisons of the correlation analyses on the pleasantness and arousal ratings between the three presentation durations.

Another way of examining differences in subjective ratings in response to IAPS pictures under different conditions of presentation duration is to compare the strength and direction of the inter-relationships of complexity, pleasantness and arousal ratings with familiarity ratings. Spearman's rank-order correlations were thus calculated (Table 1), ignoring the relationship between pleasantness and arousal due to its quadratic nature. The directions and strengths of the familiarity–pleasantness correlations were generally similar across the three presentation durations, with r_s ~.70. The familiarity-complexity correlation was not the same across the three presentation duration conditions. In the 1-s and 25-s conditions the correlation was similar in direction and strength ($r_s = -.30$), but not present in the 5-s condition. The other (non-significant) change across the three conditions of presentation duration refers to the familiarity-arousal relationship, which seems to become weaker the longer the presentation duration is, ranging from $r_s = -.52$ to $r_s = -.35$. Interestingly, the data seemed to indicate that the arousal–complexity relationship increased over time, ranging from $r_s = .18$ to $r_s = .30$. In addition, partial Spearman's rank order correlations controlling for effects of familiarity revealed that the positive complexity–arousal relationship was only significant (Bonferroni–Holm corrected) in the 5- and 25-s conditions ($r_s(87) = -.03$ for the 1-s condition, $r_s(90) = .27$ for the 5-s condition, and $r_s(87) = .21$ for the 25-s condition), whereas the relationship between complexity and pleasantness was only present in the 5-s condition ($r_s(91) = .22$) but not in the other two conditions ($r_s(88) = .10$ for the 1-s condition, and $r_s(91) = .03$ for the 25-s condition).

Four Friedman's ANOVAs were conducted on each of the four types of subjective ratings in order to investigate whether the mean ratings vary as a function of presentation duration (within-picture factor) (Fig. 2). The Friedman ANOVA on the familiarity ratings was significant, $\chi^2(2) = 109.46$, $p < .001$. A Bonferroni correction was applied and all effects are reported at a .0167 level of significance. Wilcoxon signed-rank tests showed that there was a linear increase in familiarity ratings with presentation duration. Familiarity ratings were higher in the 5-s condition than in the 1-s condition, $T = 159.5$, $r = -.54$, and also higher in the 25-s condition than in the 5-s condition, $T = 710.0$, $r = -.41$. The largest difference was observed when comparing the 1-s with the 25-s condition, $T = 61.0$, $r = -.58$. The correlations between the familiarity ratings of all pictures for the three presentation durations were as follows: $r_{s1\text{-}5} = .77$, $r_{s1\text{-}25} = .80$, and $r_{s5\text{-}25} = .84$.

For arousal ratings, a significant effect of presentation duration was found, $\chi^2(2) = 16.47$, $p < .001$. Wilcoxon signed-rank tests were used to follow up this finding and showed that arousal ratings were higher in the 5-s than the 1-s condition, $T = 1239.5$, $r = -.26$, and higher in the 25-s condition than the 1-s condition, $T = 1004$, $r = -.29$. There was no significant difference in arousal ratings between the 5-s and 25-s conditions, $T = 1799.5$, $r = -.04$. For pleasantness ratings, the Friedman test was also significant, $\chi^2(2) = 9.53$, $p = .009$. Wilcoxon signed-rank tests indicated that pleasantness ratings were higher in the 5-s condition than the 1-s condition, $T = 1235.5$, $r = -.25$. A comparison between the 25-s condition and the 5-s condition revealed lower pleasantness ratings in the 25-s condition, $T = 1507.5$, $r = -.20$. The difference in ratings between the 1-s and 25-s condition was not significant, $T = 1728$, $r = -.10$, thus indicating an inverted U-shape relationship between presentation duration and pleasantness ratings.

The Friedman ANOVA on the complexity ratings was also significant, $\chi^2(2) = 28.58$, $p < .001$. Complexity ratings were higher in the 5-s condition than in the 1-s condition, $T = 1175.5$, $r = -.24$. Complexity ratings of the 25-s condition were lower than those of the 5-s condition, $T = 389$, $r = -.49$, and also lower than those of the 1-s condition,

[1] In the two experiments, all relationships between subjective and objective variables were visually inspected in order to ensure that the distributions of the stimuli follow a linear or monotonic function.

[2] *r* refers to an effect size estimate that is derived by converting test statistics into *z*-scores and dividing by the square root of the number of total observations, $r = z/\sqrt{N}$ (Rosenthal, 1991, p. 19). The standard values of *r* for small, medium and large effect sizes are as follows: small: $r = .10$; medium: $r = .30$; large: $r = .50$.

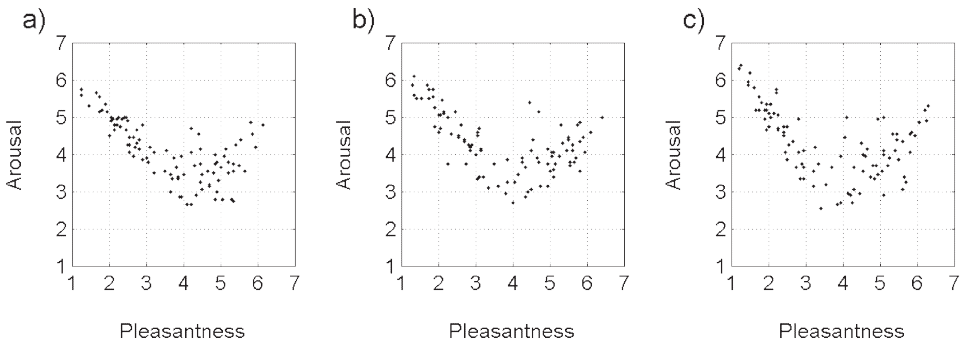

Fig. 1. a–c. Distribution of 96 IAPS pictures in the pleasantness–arousal space for three presentation durations (a — 1 s, b — 5 s, and c — 25 s). Low numbers refer to low degrees of pleasantness and arousal, respectively.

$T = 947$, $r = -.28$. Thus, complexity ratings showed an inverted U-shaped relationship across the three presentation durations. The correlations between the complexity ratings of all pictures for the three presentation durations were as follows: $r_{s1\text{-}5} = .77$, $r_{s1\text{-}25} = .80$, and $r_{s5\text{-}25} = .84$. In summary, the ratings of the four types of measures followed different types of patterns across the three presentation durations. These trends were also found when participants, and not pictures, were taken as the unit of analysis.

2.2.2. Berlyne's collative motivation model and mediation effects

In the context of Berlyne's collative motivation model (1971), Marin and Leder (2013) investigated whether arousal mediates the relationship between complexity and pleasantness and reported mediation effects for a presentation duration of 25 s. Here, we computed mediation models for each of the three conditions of presentation duration using the macro MEDCURVE developed by Hayes and Preacher (2010). This macro enables the consideration of non-linear relationships between variables. Specifically, we estimated the instantaneous indirect effect of complexity on pleasantness through arousal, conditional on the specific value of complexity, by controlling for familiarity. The results of these regressions are shown in Table 2. Results regarding the model predicting the mediator arousal indicated that complexity was only significantly related to arousal in the 5- and 25-s conditions. In a next step, the dependent variable model was computed and explained between 53 and 64% in the variance of pleasantness. The conditional indirect effects were estimated by using the sample mean and one standard deviation above and below the mean, that is, the significance of the indirect effect from complexity to pleasantness through arousal at low, moderate and high values of complexity. Confidence intervals were computed by using a resampling bootstrap technique with 5000 resamples (Hayes & Preacher, 2010). The results showed that an indirect effect was only significant in the 5- and 25-s conditions. At all three levels of complexity, an increase in complexity led to a decrease in pleasantness through an effect of complexity on arousal.

Table 1
Spearman's rank order correlations between measures of familiarity, complexity, pleasantness and arousal in response to a set of IAPS pictures for three conditions of presentation duration.

Measure		Familiarity	Complexity
Complexity	1 s	−.36* (89)	
	5 s	−.01 (92)	
	25 s	−.31* (92)	
Pleasantness	1 s	.71* (92)	−.16 (91)
	5 s	.73* (92)	.12 (94)
	25 s	.68* (93)	−.19 (93)
Arousal	1 s	−.52* (91)	.18 (90)
	5 s	−.39* (91)	.26* (93)
	25 s	−.35* (89)	.30* (89)

Note. $p^* < .05$ after Bonferroni–Holm correction. Degrees of freedom are in parentheses.

2.2.3. Changes of the relationships between subjective and objective complexity measures

To address the question of how the relationship between subjective complexity and objective complexity measures varies depending on presentation duration, a series of Spearman's rank order correlations (2-tailed) was conducted. Table 3 shows that all measures of compressed file size were significantly positively associated with subjective complexity. No significant differences across the three conditions of presentation duration were found, although the strength of the correlations unequivocally suggests an increase with presentation duration in all four compression file formats under investigation. In the 25-s condition, the correlations with the TIFF and JPEG file sizes were the strongest ($r_s \sim .50$), followed by the one with the PNG file size (all $r_s = .45$). The correlation between subjective complexity and the GIF size was the lowest ($r_s = .29$). There were no significant correlations between the different compression formats and any other type of subjective rating in all conditions of presentation duration, which replicates findings by Marin and Leder (2013).

Table 3 also shows the results of Spearman-Rho correlations (2-tailed) between different types of edge detection measures and subjective ratings of familiarity, complexity, pleasantness and arousal. All measures only correlated significantly with subjective complexity (see Marin and Leder (2013)). Similar to the findings regarding the compressed file sizes, the correlations gained in strength the longer the presentation duration was (in most cases approximately in >.10 if the difference in *r*-values between the 1-s and 25-s conditions were

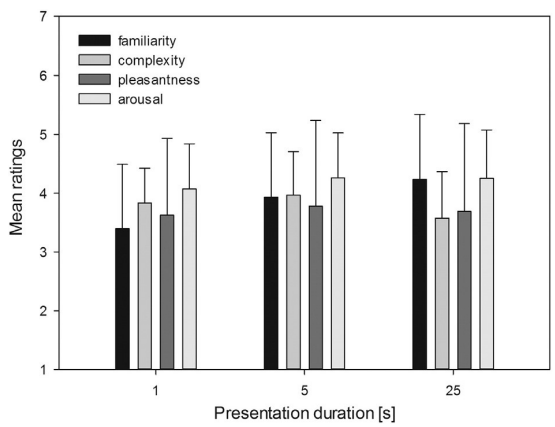

Fig. 2. Mean ratings of familiarity, complexity, pleasantness and arousal for presentation durations of 1 s, 5 s, and 25 s in response to IAPS pictures. Low numbers refer to low degrees of pleasantness, arousal and complexity, respectively. Error bars denote one standard deviation.

Table 2
Indirect effect of complexity (X) on pleasantness (Y) through arousal (M), with a quadratic relation between M and Y and familiarity as a covariate, calculated for IAPS pictures ($n_{1s} = 90$, $n_{5s} = 93$, $n_{25s} = 90$).

		Coeff	*SE*
Model predicting arousal (M)			
Constant	1 s	5.08***	.63
	5 s	4.17***	.45
	25 s	3.98***	.58
Familiarity	1 s	−.32	.07
	5 s	−.28***	.06
	25 s	−.18*	.08
Complexity (X)	1 s	.03	.13
	5 s	.30*	.09
	25 s	.29**	.11
Summary of model predicting M	1 s	$R^2 = .23$***	
	5 s	$R^2 = .25$***	
	25 s	$R^2 = .17$***	
Model predicting pleasantness (Y)			
Constant	1 s	−1.09	2.45
	5 s	−3.38	2.61
	25 s	−4.67	2.79
Familiarity	1 s	.56***	.09
	5 s	.76***	.10
	25 s	.78***	.11
Complexity (X)	1 s	.09	.16
	5 s	.39*	.13
	25 s	.14	.15
Arousal (M)	1 s	2.04	1.16
	5 s	1.83	1.28
	25 s	2.77*	1.29
Arousal squared (M*M)	1 s	−.34*	.14
	5 s	−.28	.15
	25 s	−.38*	.15
Summary of model predicting Y	1 s	$R^2 = .64$***	
	5 s	$R^2 = .62$***	
	25 s	$R^2 = .53$***	

		M_x	Θ_x	CI 95%	
Pictures of low complexity	1 s	3.24	−.03	−.20	.21
	5 s	3.21	−.12	−.28	−.03
	25 s	2.77	−.09	−.23	−.01
Pictures of moderate complexity	1 s	3.83	−.03	−.24	.19
	5 s	3.96	−.15	−.33	−.05
	25 s	3.57	−.14	−.32	−.03
Pictures of high complexity	1 s	4.42	−.03	−.27	.16
	5 s	4.72	−.19	−.41	−.06
	25 s	4.37	−.19	−.46	−.04

Note. X = independent variable, M = mediator variable, Y = dependent variable, Coeff = unstandardized coefficient, f = females, m = males, Θ_x = instantaneous indirect effect of X on Y through M at a specific value X = x, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

considered), but these differences were not significant when correlations were compared by means of a Fisher *r*-to-*z* transformation. The largest effect of presentation duration was observed for the entropy measure. Furthermore, the current results also highlight the fact that the performance of the individual measures of objective complexity remains unchanged regardless of presentation duration.

Table 3
Spearman-Rho correlations between a set of thirteen measures of objective complexity and self-reported familiarity, complexity, pleasantness and arousal in response to IAPS pictures for three presentation durations.

Measure		JPEG	GIF	PNG	TIFF	PERI-RAW	CANNY-RAW	ENTROPY	RMS-CONTR *M*	RMS-CONTR *SD*	RMS-CONTR *α*	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG *β*
Familiarity	1 s	.06	.06	.04	.05	.14	.19	.08	.07	.09	.07	.07	.03	.07
	5 s	.13	.11	.10	.12	.22	.18	.10	.15	.14	.15	.08	.07	.09
	25 s	−.002	.08	.03	.04	.14	.12	.11	.04	.04	.03	−.01	−.03	−.004
Complexity	1 s	.40*	.20	.33	.38*	.29	.12	.06	.38*	.34*	.41*	.22	.26	.24
	5 s	.46*	.24	.38*	.46*	.39*	.19	.23	.55*	.47*	.55*	.29	.33*	.32
	25 s	.50*	.29	.45*	.52*	.42*	.29	.31	.57*	.39*	.54*	.33*	.32	.34*
Pleasantness	1 s	.11	.09	.09	.11	.10	.15	.03	.09	.17	.14	.06	.09	.07
	5 s	.12	.12	.12	.15	.12	.17	.06	.11	.17	.17	.08	.10	.09
	25 s	.08	.09	.09	.11	.07	.15	−.01	.06	.18	.14	.07	.11	.08
Arousal	1 s	.01	.01	.003	−.01	−.02	−.01	.04	.03	−.07	−.04	−.11	−.16	−.12
	5 s	.09	.11	.10	.08	.09	.05	.08	.11	−.04	.02	−.01	.001	.03
	25 s	.04	.04	.001	.01	.04	.002	.15	.07	−.06	.02	−.10	−.15	−.09

Note. $p^* < .05$ after Bonferroni–Holm correction, PERI = perimeter, CONTR = contrast,CONG = congruency, RMS = root mean square.

In summary, we observed effects of presentation duration on most analyses involving subjective ratings, such as for the comparison of the emotion spaces and of the averaged ratings of familiarity, complexity, pleasantness and arousal, and to a lesser degree, for the comparison of the inter-relationships of these variables across time. The mediation analysis indicated that only in the 5- and 25-s conditions mediation effects of arousal were present. The correlation strengths for objective measures with subjective complexity were not significantly affected by presentation duration, although for certain measures an increase in correlation strength over time was observed.

2.3. Discussion

The observed general similarity of the emotion spaces (i.e., the inverted U-shape) across the 1-s, 5-s and 25-s conditions is in line with previous investigations of the time course of the IAPS emotion space for short presentation durations between 25 ms and 6 s (Codispoti et al., 2009). This may suggest that responses to affective environmental scenes do not vary much over time, possibly due to the pictures' motivational relevance for humans. However, our analyses also indicate that changes in affective responses may occur for presentation durations above 1 s. To be specific, the emotion spaces of the 5-s and 25-s conditions were more similar to each other than to the emotion space of the 1-s condition. One possible explanation may relate to higher-order cognitive influences (e.g., associations and memories triggered by the affective contents) on affective ratings that may become more influential under longer viewing times and moderate the initial affective responses after picture onset.

The observation of changes in arousal and pleasantness ratings for presentation durations between 1 and 25 s is not in line with results reported by Codispoti et al. (2009), who demonstrated that in short exposure times of 25 ms to 6 s pleasantness and arousal ratings as well as physiological indicators in response to IAPS pictures did not change considerably after early processing stages. The authors interpreted their results by arguing that people's emotional engagement mostly relies on the affective information transmitted by the semantic contents of these environmental scenes. A closer inspection of their data revealed that arousal ratings were the highest in the 5-s and 6-s conditions. Unfortunately, Codispoti et al. (2009) only had 500-ms and 1.5 s conditions but no 1-s condition and did not offer any exact statistical results on differences between these conditions, which makes a direct comparison impossible. However, Codispoti et al.'s findings suggest that arousal ratings were lower in the 1.5-s condition than in the 5-s condition, which would be in line with our current finding of higher arousal ratings for longer exposure durations than in the 1-s condition.

With regard to pleasantness ratings, our findings do not corroborate the results by Codispoti et al. (2009) who reported that changes in exposure durations larger than 250 ms and up to 6 s did not affect pleasantness ratings. We observed significant higher ratings in the 5-s condition than in the 1-s and 25-s conditions, a result that does not support theories linking processing fluency with affective judgments. According to fluency theory, longer presentation durations should be accompanied by higher preference ratings due to ease of perceptual processing (Forster, Leder, & Ansorge, 2013; Reber, Schwarz, & Winkielman, 2004), which in the current could only explain the difference in pleasantness ratings between the 1-s and 5-s conditions but not the lower pleasantness ratings in the 25-s condition. Furthermore, familiarity has also been discussed in relation to fluency by suggesting that high fluency is positive because the stimulus appears familiar (Winkielman, Schwarz, Fazendeiro, & Reber, 2003). Correlations between familiarity and pleasantness were of a moderate to large effect size and did not vary across the three presentation durations. Interestingly, the average familiarity ratings in the 25-s condition were the highest among the three presentation durations, but average ratings of pleasantness were similar to the 5-s condition. However, the ease of perceptual and conceptual processing presumably plays a minor role in determining pleasantness under long viewing conditions.

One possible explanation for higher familiarity ratings with increasing exposure duration may relate to the fact that longer exposure durations enhance the ease of memory retrieval processes necessary to judge familiarity with the picture content. In other words, the longer the presentation duration, the "more correct" may be the familiarity judgment in response to the semantic contents of the environmental scene. Although inter-rater reliability of familiarity ratings was similar across the three presentation duration conditions, more correct judgments may indeed have occurred under longer presentation durations. This hypothesis may be tested with controlled memory paradigms in the future. Also, as suggested by Monin (2003), changes in presentation times that change liking or affective evaluations as in the warm glow heuristic might increase false feelings of familiarity.

The results regarding the subjective experience of complexity showed that the highest complexity ratings were observed for the 5-s condition and the lowest for the 25-s condition. It can be hypothesized, according to global precedence (Navon, 1977), that in the 1-s condition mostly bottom-up processes determine complexity ratings and that an exposure time of 1 s is sufficient to perceive the general complexity of an image (global processing). In the 5-s condition, however, participants may have had more time to explore the image thoroughly and to discover details that may have led to higher complexity ratings (local processing). In the 25-s condition, complexity ratings were the lowest, which may be explained by top-down phenomena. Since participants also reported the highest familiarity in the 25-s condition, it is likely that this high degree of felt familiarity was associated with a reduction of uncertainty, which in turn may have led to lower complexity judgments than in the 1-s condition. Future experiments may thus systematically explore the relationship between familiarity, complexity, and mechanisms of global/local processing in response to environmental scenes across different exposure durations.

Mediation analysis revealed an indirect effect of complexity on pleasantness through arousal in the 5- and 25-s conditions but not in the 1-s condition. Notably, the correlations between complexity and arousal as well as the one between complexity and pleasantness were not significant in the 1-s condition. This may suggest that subjective reports with regard to affective environmental scenes may only support Berlyne's psychobiological model (1971) when stimuli are presented for more than 1 s. This hypothesis could be further tested by employing presentation durations that are below 1 s, to confirm that mediation is not present, and between the range of 1 and 5 s, to examine which minimum exposure duration is necessary in order to observe mediation at the level of subjective experience.

Our current results support and extend earlier reports of a positive correlation between compressed file size and subjective complexity reported for different types of stimuli and presentation durations (Cavalcante et al., 2014; Donderi, 2006b; Forsythe et al., 2008; Forsythe et al., 2011). To be specific, we showed that correlation strengths did not change significantly across the three presentation durations, although indications of stronger correlations were present for longer presentation durations. One possible explanation for the increase in the correlation strength over time (>.10 on average) may relate to a decrease in subjective uncertainty in the rating process. Although it is a well-established finding that object classification only requires between 40 and 67 ms (Fei-Fei, Iyer, Koch, & Perona, 2007), judgments of environmental scenes clearly go beyond mere object recognition (e.g., Greene & Oliva, 2009), and further, may be more prone to individual differences in short presentation times. Therefore, subjective complexity judgments may still change over time and become more like objectively measurable complexity as expressed by a stronger correlation with objective complexity in the 25-s condition. Furthermore, this finding highlights the fact that the pattern of results, that is, the order of the performance of the individual measures of objective complexity form high to low, remains unchanged when presentation time is kept constant.

3. Experiment 2

3.1. Method

3.1.1. Participants

Sixty female German-speaking undergraduate students participated in the experiment in return for course credit. Most of these students were enrolled in a psychology program and none of them was pursing an art history degree. Participants did not take part in any other experiment of the current study, and they all had normal or corrected-to-normal visual acuity. Twenty females were tested for each of the three experimental conditions varying in stimulus presentation time. The mean age of these three participant groups (1-s condition: 23.6 ± 3.3 years; 5-s condition: 23.3 ± 4.5 years; 25-s condition: 22.0 ± 1.5 years) did not significantly differ, $H(2) = 2.46$, $p = .292$. The 20 participants of the 25-s condition are those that participated in Experiment 2 reported in Marin and Leder (2013).

3.1.2. Materials

Ninety-six digital high-quality representations of colored oil and acrylic paintings served as stimuli. As described in Marin and Leder (2013), these representational paintings included different styles from the end of the 18th to the beginning of the 20th century. Their semantic contents was similar to the one of the IAPS pictures used in Experiment 1, comprising human beings in different situations of everyday life (excluding erotic scenes), landscapes, animals and plants. These stimuli were pre-selected to fall into one of the four quadrants spanned by arousal and pleasantness (Russell, 1980). Half of the paintings in each hedonic category ($n = 24$) were pre-selected to depict a figure-ground composition and half a complex scene. The paintings were all in landscape format and saved in the same size (1024 × 768) and resolution (72 dpi) in JPEG format (maximum quality) as the IAPS pictures used in Experiment 1.

Two self-developed questionnaires were administered, one asking specific questions referring to the experiment and one probing general interest in arts and expertise in visual arts. The former questionnaire comprised two specific questions about the paintings, namely whether participants recognized the styles of the paintings and could name the century of their creation. Participants were also asked to indicate the degree of their general liking for the paintings on a seven-point Likert scale ("not at all"…"very much"). Moreover, participants reported the general level of difficulty associated with giving ratings of complexity on a seven-point Likert scale ("very easy"…"very difficult"). Last, they estimated the percentage of paintings they had seen at least once before the experiment (ranging from 0% to 100%). The questionnaire on interest in arts and expertise in visual arts comprised three parts. A set of 12 questions was used to examine participants' general interest in visual arts (on a seven-point Likert scale ranging from "fully disagree" to "fully agree"), not considering specific styles or visual art forms. For instance, participants had to rate whether they regularly visit art exhibitions in their spare time or whether they like to read about art. In the subsequent part, participants reported whether they were familiar with the names of a range of artists, their nationality as well as with the associated style. In the final section, participants were shown reproductions of six modern paintings and again had to indicate their familiarity with the paintings and to name the artist and the style if possible.

3.1.3. Procedure

The procedure of Experiment 2 did generally not differ from the one described in Experiment 1 (see Marin and Leder (2013)) in order to make a meaningful comparison of results feasible. One difference with regard to Experiment 1 concerns the instructions for familiarity ratings, which were as follows: "Please rate your familiarity with the painting." In addition, after participants had finished the experiment, they filled a self-developed one-page questionnaire on aspects of the experiment followed by a four-page questionnaire, assessing general art expertise and interest in visual arts. These questionnaires extended the length of the experimental sessions for approximately ten minutes compared to Experiment 1. The same type of compression formats and edge detection measures were applied to the paintings as described in Experiment 1 as in Marin and Leder (2013).

3.2. Results

The results are presented in the same way as in Experiment 1.

3.2.1. Changes of subjective ratings over time

There was no significant difference in self-reported interest in visual arts among the three presentation duration conditions, $F(2, 57) = 1.00$, $p = .374$. Further inspection of the questionnaire data suggested that participants were not art experts. Normality checks of the questionnaire data on participants' liking for the paintings used in the experiment, the level of difficulty of giving complexity judgments as well as the frequency of having seen the paintings prior to the experiment revealed that all except two out of 12 Kolmogorov-Smirnov normality tests were significant (all $ps < .14$). Kruskal–Wallis tests showed that the three groups gave similar liking ratings, $H(2) = .99$, $p = .611$, and ratings of the difficulty of complexity judgments, $H(2) = 1.99$, $p = .370$. However, groups differed significantly with regard to the self-reported frequency of having seen the paintings prior to the experiment, $H(2) = 10.74$, $p = .005$. Further analyses showed that participants in the 25-s condition reported to have seen fewer paintings (0–5%) prior to the experiment than those in the 1-s condition (5–10%), Mann–Whitney $U = 87.50$, $p = .002$, $r = -.42$ (Bonferroni corrected). This finding was not regarded as particularly critical since familiarity ratings were considered, whenever possible, in the main analyses. No other significant differences between groups were observed.

As in Experiment 1, inter-rater reliability was assessed by calculating the average intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency (McGraw & Wong, 1996; Shrout & Fleiss, 1979). We observed a very high inter-rater reliability for all four types of ratings in all three conditions of presentation duration. For the presentation duration of 1 s, the results were as follows: familiarity (ICC(2, 20) = .82, 95% confidence interval (CI) [.76, .87]), complexity (ICC(2, 20) = .87, 95% CI [.82, .90]), pleasantness (ICC(2, 20) = .95, 95% CI [.93, .96]), and arousal (ICC(2, 20) = .79, 95% CI [.72, .85]). Similar results were obtained for the presentation duration of 5 s: familiarity (ICC(2, 20) = .94, 95% CI [.92, .96]), complexity (ICC(2, 20) = .90, 95% CI [.87, .93]), pleasantness (ICC(2, 20) = .94, 95% CI [.92, .96]), and arousal (ICC(2, 20) = .83, 95% CI [.77, .87]). Inter-rater reliability was also considered as sufficiently high for the 25-s condition: familiarity (ICC(2, 20) = .82, 95% CI [.76, .87]), complexity (ICC(2, 20) = .87, 95% CI [.83, .90]), pleasantness (ICC(2, 20) = .94, 95% CI [.92, .95]), and arousal (ICC(2, 20) = .84, 95% CI [.79, .88]). As in Experiment 1, the inter-rater reliability of each variable was similar across the three presentation durations, justifying the comparison of the pictures' ratings across the three conditions.

Data screening of the variables referring to different compression formats for outliers revealed between zero and six outliers per variable. Outliers were defined by a file size of ±2 *SD* from the mean size and removed. A series of Kolmogorov–Smirnov normality tests indicated that the variables of the compression file formats were normally distributed, as was the case for the measure of Canny edge detection. The measure of perimeter detection and the RMS α-measure were positively skewed and transformed by using a logarithmic transformation. The data regarding the entropy of the paintings was negatively skewed and reversed prior to the application of a logarithmic transformation. Subjective ratings of all pictures in each condition of presentation duration were averaged across participants. After the deletion of outliers, Kolmogorov-Smirnov tests showed that only two out of 12 variables regarding self-reported familiarity, complexity, pleasantness and arousal in three conditions of presentation duration significantly deviated

from a normal distribution (arousal and pleasantness, 5-s condition, $ps = .03$). Based on further visual inspection of the data, it was decided to perform parametric analyses.

Fig. 3a–c shows that the distributions of the arousal and pleasantness ratings in the three conditions of presentation duration are following quadratic relationships, which is in line with the findings of Experiment 1. However, the strength of the quadratic relationship is weaker for paintings than for environmental scenes ($.62 \le R^2 \le .69$). For the 1-s condition, a curve fitting analysis revealed the following results: $R^2 = .42$, $F(2, 89) = 32.73$, $p < .001$, $y = -1.74x + .19x^2 + 7.57$. The quadratic relationship was stronger in the 5-s condition, $R^2 = .61$, $F(2, 92) = 70.87$, $p < .001$, $y = -2.44x + .25x^2 + 9.07$, whereas in the 25-s condition, the quadratic relationship was similar to the one observed in the 1-s condition: $R^2 = .38$, $F(2, 89) = 27.37$, $p < .001$, $y = -2.18x + .25x^2 + 8.33$. The Pearson product–moment correlations for arousal ratings were as follows: $r(90) = .74$ for $arousal_{1s}$ – $arousal_{5s}$, $r(88) = .71$ for $arousal_{1s}$ – $arousal_{25s}$, and $r(90) = .77$ for $arousal_{5s}$ – $arousal_{25s}$. Similar to results in Experiment 1, somewhat higher correlations were observed for pleasantness ratings across the three conditions of presentation duration: $r(94) = .95$ for $pleasantness_{1s}$ – $pleasantness_{5s}$, $r(94) = .95$ for $pleasantness_{1s}$ – $pleasantness_{25s}$, and $r(94) = .96$ for $pleasantness_{5s}$ – $pleasantness_{25s}$.

In order to make a more comprehensive, global comparison between the different emotion spaces possible, three Procrustes analyses (Bookstein, 1997; Kendall, 1984) based on the averaged pleasantness and arousal ratings were performed in Matlab. Two common measures reported in the examination of the similarity of the two shapes are a standardized dissimilarity measure d (ranging between 0 and 1 with lower numbers indicating higher similarity) and a scaling measure b that refers to the size similarity of the two shapes. In general, the emotion spaces differed to a larger degree from each other ($.176 \le d \le .196$) than those of environmental scenes in Experiment 1 ($.085 \le d \le .118$). Specifically, the comparison between the 1-s and 5-s conditions yielded the following results: $d = .196$ and $b = .782$. The comparison between the 1-s and 25-s conditions ($d = .189$, $b = .91$) suggests that the shapes are more similar to each other than those of the 1-s and 5-s conditions. As in Experiment 1, the final Procrustes analysis compared the emotion spaces of the 5-s and 25-s conditions, and the results indicated the highest similarity in shape ($d = .176$) and size ($b = 1.051$) among the three comparisons.

For each of the three conditions of presentation duration, Table 4 shows the nature of the relationships between the subjectively reported variables of familiarity, complexity, pleasantness and arousal. With the exception of the familiarity-arousal relationship, the strength of the Pearson product–moment correlations seemed to depend on the three conditions of presentation duration: the shorter the presentation duration, the weaker the respective correlation. The largest increase in strength was observed for the familiarity-complexity relationship ($\Delta = .3$), followed by the complexity-arousal ($\Delta = .25$) and complexity-pleasantness ($\Delta = .2$) relationships. Partial correlations, controlling for familiarity and computed for each presentation duration, revealed that the relationship between complexity and arousal was lower in the 1-s condition than in the two longer presentation duration conditions, $r_{1s}(85) = .37$, $r_{5s}(85) = .52$ and $r_{25s}(86) = .49$. The pleasantness-complexity relationship was not significant in any of the three time conditions, $r_{1s}(87) = -.12$ and $r_{5s}(85) = -.15$, $r_{25s}(89) = -.02$.

Table 4
Pearson product–moment correlations between ratings of familiarity, complexity, pleasantness and arousal in response to a set of representational paintings in three conditions of presentation duration.

Measure		Familiarity	Complexity
Complexity	1 s	.03 (88)	
	5 s	−.09 (87)	
	25 s	−.32* (90)	
Pleasantness	1 s	.69* (93)	−.07 (89)
	5 s	.77* (91)	−.17 (90)
	25 s	.80* (93)	−.27* (91)
Arousal	1 s	−.49* (89)	.31* (86)
	5 s	−.49* (90)	.49* (89)
	25 s	−.44* (89)	.56* (87)

Note. $p^* < .05$ after Bonferroni–Holm correction. Degrees of freedom are in parentheses.

Next, we assessed whether the mean ratings of each of the subjective variables changed over the three conditions of presentation duration. Four repeated-measures ANOVAs with presentation duration as a within-picture factor on the subjective variables (Fig. 4) revealed significant effects for familiarity, $F(1.86, 170.93) = 12.57$, $p < .001$, $\eta^2_p = .12$, complexity, $F(2, 176) = 22.87$, $p < .001$, $\eta^2 = .21$, arousal, $F(2, 178) = 13.36$, $p < .001$, $\eta^2 = .13$, as well as for pleasantness ratings, $F(2, 190) = 11.49$, $p < .001$, $\eta^2 = .11$. Pairwise comparisons indicated that familiarity ratings in the presentation duration of 1 s ($M = 2.69$, $SEM = .07$) were higher ($p = .007$) than in the 5-s condition ($M = 2.56$, $SEM = .06$). Familiarity ratings were also higher in the 25-s condition ($M = 2.81$, $SEM = .06$) than in the 5-s condition, $p < .001$. Significant correlations were observed for familiarity ratings across the three conditions of presentation duration: $r(91) = .80$ for $familiarity_{1s}$ – $familiarity_{5s}$, $r(93) = .69$ for $familiarity_{1s}$ – $familiarity_{25s}$, and $r(91) = .73$ for $familiarity_{5s}$ – $familiarity_{25s}$. For complexity ratings, there was a significant difference between the 1-s and 5-s conditions ($p < .001$), with higher complexity ratings for the 5-s condition ($M = 4.28$, $SEM = .08$) than in the 1-s condition ($M = 4.04$, $SEM = .07$). Furthermore, there was a significant difference between the 5-s and 25-s conditions ($p < .001$), with lower mean ratings for the 25-s condition ($M = 4.1$, $SD = .06$). The inter-correlations of complexity ratings for different presentation durations were as follows: $r(87) = .82$ for $complexity_{1s}$ – $complexity_{5s}$, $r(87) = .81$ for $complexity_{1s}$ – $complexity_{25s}$, and $r(90) = .86$ for $complexity_{5s}$ – $complexity_{25s}$. With regard to arousal ratings, mean ratings in the 1-s condition ($M = 3.91$,

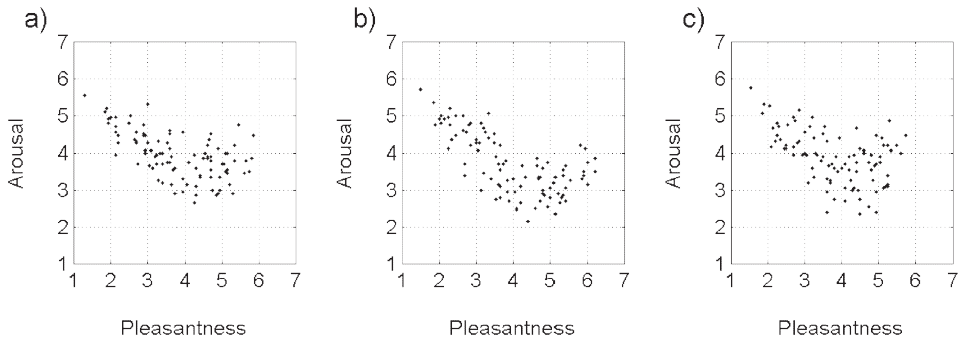

Fig. 3. a–c. Distribution of 96 representational paintings in the pleasantness-arousal space for three presentation durations (a — 1 s, b — 5 s, c — 25 s). Low numbers refer to low degrees of pleasantness and arousal.

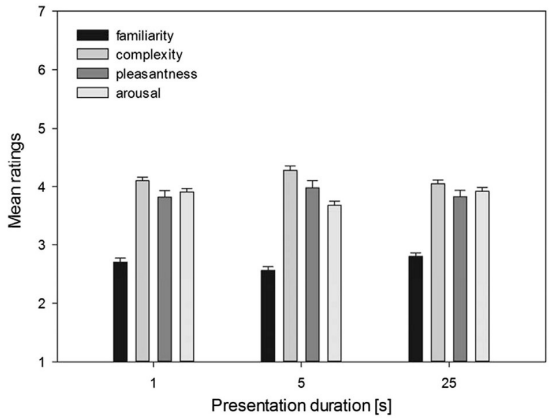

Fig. 4. Mean ratings of familiarity, complexity, pleasantness and arousal for presentation durations of 1 s, 5 s, and 25 s in response to representational paintings. Low numbers refer to low degrees of pleasantness, arousal and complexity, respectively. Error bars denote standard errors of the means.

SEM = .06) were higher ($p < .001$) than in the 5-s condition ($M = 3.67$, *SEM* = .08). The ratings in the 25-s condition ($M = 3.92$, *SEM* = .07) were also higher than those in the 5-s condition, $p < .001$. For pleasantness ratings, average ratings were lower ($p < .001$) in the 1-s condition ($M = 3.82$, *SEM* = .11) than in the 5-s condition ($M = 3.98$, *SEM* = .11). Ratings in the 25-s condition ($M = 3.83$, *SEM* = .11) were lower ($p < .001$) than those in the 5-s condition. This pattern of results was also present when participants, and not pictures, were taken as a unit of analysis.

3.2.2. Berlyne's collative motivation model and mediation effects

Marin and Leder (2013) investigated whether arousal mediates the relationship between complexity and pleasantness and reported mediation effects for a presentation duration of 25 s for paintings of moderate and high complexity in females. As in Experiment 1, we also computed mediation models for each of the three conditions of presentation duration. The results of these regressions are shown in Table 5. Results regarding the model predicting the mediator arousal indicated that complexity was significantly related to arousal in all conditions. In a next step, the dependent variable model was computed and explained between 57 and 73% in the variance of pleasantness. The results showed that an indirect effect was significant for paintings of moderate and high

Table 5
Indirect effect of complexity (X) on pleasantness (Y) through arousal (M), with a quadratic relation between M and Y and familiarity as a covariate, calculated for representational paintings ($n_{1s} = 87$, $n_{5s} = 88$, $n_{25s} = 88$).

		Coeff	*SE*
Model predicting arousal (M)			
Constant	1 s	3.82***	.40
	5 s	3.50***	.48
	25 s	2.88***	.54
Familiarity	1 s	−.47***	.08
	5 s	−.72***	.11
	25 s	−.34***	.10
Complexity (X)	1 s	.34*	.08
	5 s	.47***	.08
	25 s	.48***	.09
Summary of model predicting M	1 s	$R^2 = .37$***	
	5 s	$R^2 = .49$***	
	25 s	$R^2 = .40$***	
Model predicting pleasantness (Y)			
Constant	1 s	−2.95	3.31
	5 s	−2.21	1.59
	25 s	−2.75	2.03
Familiarity	1 s	1.0***	.15
	5 s	1.12***	.15
	25 s	1.24***	.12
Complexity (X)	1 s	−.02	.15
	5 s	.09	.11
	25 s	.22	.12
Arousal (M)	1 s	2.72	1.80
	5 s	2.24*	.85
	25 s	1.59	1.02
Arousal squared (M*M)	1 s	−.41	.22
	5 s	−.38**	.11
	25 s	−.26*	.12
Summary of model predicting Y	1 s	$R^2 = .57$***	
	5 s	$R^2 = .73$***	
	25 s	$R^2 = .69$***	

		M_x	Θ_x	CI 95%	
Pictures of low complexity	1 s	3.49	−.10	−.26	.03
	5 s	3.54	−.12	−.28	.03
	25 s	3.47	−.13	−.30	.03
Pictures of moderate complexity	1 s	4.08	−.15	−.33	−.03
	5 s	4.29	−.24	−.41	−.12
	25 s	4.08	−.20	−.37	−.07
Pictures of high complexity	1 s	4.67	−.21	−.43	−.05
	5 s	5.04	−.36	−.62	−.19
	25 s	4.69	−.28	−.49	−.12

Note. X = independent variable, M = mediator variable, Y = dependent variable, Coeff = unstandardized coefficient, f = females, m = males, Θ_x = instantaneous indirect effect of X on Y through M at a specific value X = x, *$p < .05$, **$p < .01$, ***$p < .001$.

complexity in all three presentation duration conditions. An increase in complexity led to a decrease in pleasantness through an effect of complexity on arousal.

3.2.3. Changes of the relationships between subjective and objective complexity measures

Pearson product–moment correlations were computed to examine how the relationships between subjective and objective complexity as measured by 13 different measures (see Experiment 1) changed over presentation durations. In general, the pattern of results corresponded to the one reported in Experiment 2 in Marin and Leder (2013). Table 6 shows that none of the compressed file sizes correlated significantly with any of the four subjective variables. However, significant correlations with subjective complexity were observed for the phase congruency *SD* measure, and somewhat weaker positive correlations were present for the RMS contrast measures as well as negative correlations for the Canny measure.

The current data suggests that some correlations were (non-significantly) stronger in the 1-s and 5-s conditions than in the 25-s condition. This holds true for measures of perimeter detection, entropy, and RMS contrast, which stands in contrast to findings related to IAPS picture for which the reverse was observed, i.e. stronger correlations between objective and subjective complexity in longer presentation durations. Correlations of the Canny algorithm and the three measures of phase congruency did not vary among the three time conditions, indicating that some measures are more robust against effects of presentation duration than others.

RMS contrast measures of objective complexity also correlated significantly with arousal in the 5-s condition and to a lesser degree in the other time conditions. Some trends regarding effects of presentation duration on correlations with arousal were noted for other measures of objective complexity: For example, the correlation of phase congruency measures seemed to be stronger in the 25-s condition than in the shorter conditions. The current data clearly indicates that the observed trends over different presentation durations may depend on the specific measure of objective complexity under consideration.

To summarize, we found effects of presentation duration regarding the emotion spaces spanned by the paintings, the averaged ratings of familiarity, complexity, pleasantness and arousal across paintings, as well as for the inter-relationships of these variables. Regarding the nature of the specific effects, the pattern of results differed somewhat from the one observed in environmental scenes. The results of the mediation analyses in the context of Berlyne's theory (1971) revealed the presence of an indirect effect even in the 1-s condition, but only for paintings of moderate and high complexity. There were no significant effects of presentation duration regarding the relationship between objective measures of complexity and subjective ratings.

3.3. Discussion

The comparison of the 2-dimensional emotion spaces across the three presentation durations by means of Procrustes analysis suggested a higher similarity of spaces based on ratings obtained for 5 and 25 s. Experiment 1 revealed a similar, perhaps even more convincing result, which makes it likely that emotion spaces of visual stimuli for presentation durations of 5 s and longer are more stable and differ from those based on ratings of 1 s and below. Such a hypothesis could be tested further by applying different stimulus sets and a larger set of presentation durations. Interestingly, arousal ratings across the three presentation durations were less correlated with each other than those observed for environmental scenes (Exp. 1), whereas the correlations between pleasantness ratings were similar for both types of visual stimuli. This may be due to the fact that arousal ratings were more strongly influenced by presentation duration in the case of representational paintings. However, the slightly weaker inter-rater agreement of arousal ratings observed for paintings may account for the lower correlations between arousal ratings across the three presentation durations.

The current findings on the emotion spaces of representational paintings further showed that unpleasant representational paintings received higher arousal ratings than pleasant ones, which may reflect the presence of a negativity bias (e.g., Lang et al., 2005; Rozin & Royzman, 2001; Vaish, Grossmann, & Woodward, 2008) in the processing of esthetic stimuli, extending reports by Marin and Leder (2013). In general, negative emotional responses to visual arts are rarely investigated (but see Leder, Gerger, Brieber, & Schwarz, 2014; Silvia, 2009; Silvia & Brown, 2007), thus it needs to be seen whether this finding can be replicated for abstract art works and is not merely due to the representational nature of the current stimuli.

A comparison of the inter-relationships of the subjective ratings between the current set of representational paintings (Table 4) and the environmental scenes used in Experiment 1 (Table 1) revealed several differences in terms of strength and changes over time. Whereas in paintings the strength of the relationship between familiarity and complexity as well as between familiarity and pleasantness seemed to increase with presentation duration, this trend was not clearly present in environmental scenes. Furthermore, the strength of the correlation between familiarity and arousal did not vary according to presentation duration for paintings, but got slightly weaker with longer presentation durations in environmental scenes. Here, it is important to point out that familiarity was not rated in the same way in both experiments (familiarity with semantic contents in environmental scenes versus specific painting). The complexity-pleasantness relationship seemed to gain in strength in paintings, which was not the case in environmental scenes. Controlling for familiarity effects, the complexity-pleasantness relationship decreased in environmental scenes, but no such trend

Table 6
Pearson product–moment correlations between a set of thirteen measures of objective complexity and self-reported familiarity, complexity, pleasantness and arousal in response to representational paintings for three presentation durations.

Measure		JPEG	GIF	PNG	TIFF	LG PERI-RAW	CANNY-RAW	LG ENTROPY REV	RMS-CONTR *M*	RMS-CONTR *SD*	LG RMS-CONTR *α*	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG *β*
Familiarity	1 s	.19	.06	.22	.18	.07	.09	−.16	.12	−.008	−.009	.001	.003	.003
	5 s	.19	−.06	.18	.14	.12	.13	−.11	.20	.06	.09	.11	.10	.10
	25 s	.09	−.06	.10	.07	−.03	.14	−.10	.03	−.13	−.11	−.08	−.08	−.09
Complexity	1 s	.19	.15	.13	.19	.23	−.23	−.17	.28	.36	.32	.27	.39*	.33
	5 s	.06	.13	.04	.08	.21	−.24	−.17	.21	.31	.29	.30	.45*	.37*
	25 s	.04	.13	−.04	.01	.12	−.24	−.01	.10	.24	.21	.26	.40*	.34
Pleasantness	1 s	.12	−.07	.09	.10	.01	.20	−.07	−.01	−.22	−.18	−.07	−.11	−.09
	5 s	.11	−.13	.09	.07	.001	.18	−.008	−.04	−.18	−.17	−.03	−.05	−.05
	25 s	.19	−.06	.14	.15	.004	.21	−.07	−.03	−.23	−.20	−.09	−.12	−.11
Arousal	1 s	.04	.15	.01	.10	.08	−.08	−.10	.16	.29	.31	.07	.12	.08
	5 s	−.03	.08	−.02	.01	.05	−.22	−.17	−.18	.35*	.34*	.09	.22	.15
	25 s	−.10	.05	−.19	−.08	−.03	−.22	−.09	.09	.28	.24	.20	.27	.24

Note. $p^* < .05$ after Bonferroni-Holm correction, PERI = perimeter, CONTR = contrast, CONG = congruency, RMS = root mean square, LG = logarithm.

was seen with regard to paintings anymore. In both stimuli sets, the strength of the complexity-arousal relationship, which was stronger in paintings, increased with presentation duration. After controlling for familiarity, this pattern of results was only present in paintings. In summary, the current data suggest that trends of changes in correlational strength seemed to be more evident with regard to representational paintings than environmental scenes and thus be indicative of esthetic experiences, which dynamically develop over time (Leder et al., 2004).

Complexity and pleasantness ratings changed in a similar way in environmental scenes and representational paintings across the three conditions of presentation duration, suggesting an inverted U-shape relationship (see Fig. 5 for a summary of these relationships). This may indicate a general mechanism that holds true independently of the type of visual input. Whether this pattern of results will also be present for other types of visual stimuli, such as abstract patterns or abstract paintings, remains an open question. Cupchik and Berlyne (1979) investigated presentation durations between 50 ms and 5 s and reported an increase in pleasantness, which corresponds to our present finding for this time span.

Differences compared to results reported in Experiment 1 concern familiarity ratings, which increased linearly with presentation duration in environmental scenes, but showed a U-shape relationship in paintings. However, effects of presentation duration on familiarity should be interpreted with caution since ratings were obtained in a between-subjects design. In environmental scenes, arousal was higher in the two longer presentation duration conditions than in the 1-s condition, which was not the case in paintings. This latter finding, especially the decrease of arousal between the 1-s and 5-s conditions, is in line with Cupchik and Berlyne (1979), who reported lower ratings on the relaxed-tensed scale for similar exposure durations.

Results of the mediation analyses in the context of Berlyne's theory (1971) revealed significant mediation for moderate and complex paintings but not for simple ones in all conditions of presentation duration. These findings stand in contrast to the results of Experiment 1, where indirect effects were only significant in the two longer presentation duration conditions for all three levels of picture complexity. It is likely that mediation and its change over time are dependent on the type of visual stimulus and, as indicated by the interaction, the variation of complexity within the stimulus set.

With regard to measures of objective complexity, the current findings extend the results reported by Marin and Leder (2013), who argued that the background of simple figure-ground compositions cannot be compressed to a large extent due to the natural variability of the brush marks present in most of these paintings, therefore compressed file size cannot be a correlate of subjective complexity. The only objective measures yielding significant correlations with subjective complexity were based on phase congruency. Measures of RMS contrast and the Canny algorithm revealed somewhat weaker correlations with subjective complexity. Effects of presentation duration on these measures were ambiguous. For instance, the correlations for the RMS contrast measures seemed to decrease with longer presentation durations, whereas the correlations with the phase congruency measures did not seem to vary over time. Altogether, these results further support the claim that not only the performance of objective measures (Chikhman et al., 2012; Marin & Leder, 2013), but also their power to predict changes over time, may depend on the type of visual stimulus under investigation.

In order to further explore the effects regarding dynamic changes in perceived complexity, Experiment 3 explores the time course of the multidimensionality of visual complexity in comparison to general complexity ratings obtained in Experiments 1 and 2. For this purpose, the very same stimulus sets and objective measures were used.

4. Experiment 3

4.1. Method

4.1.1. Participants

One hundred nine German-speaking female psychology undergraduates participated in exchange for course credit. All participants had normal or corrected-to-normal visual acuity. Fifty-seven participants were tested in the conditions involving judgements of environmental scenes (IAPS pictures). The mean age of participants ($n = 27$) in the 1-s condition was 21.3 ± 3.3 years (range = 18–33 years), and the mean age of those ($n = 31$) in the 5-s condition was 21.8 ± 2.5 years (range = 18–30 years). The two groups did not differ in their mean age, Mann–Whitney $U = 474.5$, $p = .261$. Fifty-two participants of the two conditions evaluating paintings did also not differ in their mean age, Mann–Whitney $U = 232$, $p = .051$. The mean age of participants ($n = 26$) in the 1-s condition was 24.0 ± 5.0 years (range = 18–34 years), and the mean age of those ($n = 26$) in the 5-s condition 21.3 ± 2.9 years (range = 18–31 years). Further analysis revealed that the mean age of the four groups did not differ, $H(3) = 5.81$, $p = .121$. These participants did not take part in any other experiment of this study. In the analysis, general complexity ratings of IAPS pictures

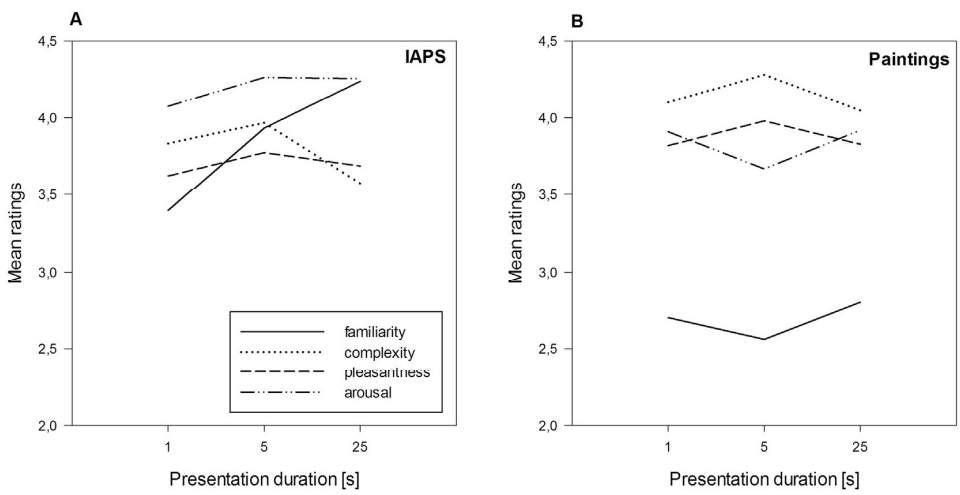

Fig. 5. Mean ratings of familiarity, complexity, pleasantness and arousal for presentation durations of 1 s, 5 s, and 25 s in response to a) environmental scenes and b) representational paintings. Low values refer to low degrees of pleasantness, arousal and complexity, respectively.

(Experiment 1) and paintings (Experiment 2) for the 1-s and 5-s condition were also considered.

4.1.2. Materials

The stimuli (i.e., IAPS pictures and representational paintings) were identical to those used in Experiments 1 and 2. Six questions referring to the experiment asked the participants how they liked the images in general, how many of the images they had previously seen, whether they felt feelings while viewing the images, and how difficult it was to give ratings. In addition, the self-developed questionnaire on visual art interest was used (see Experiment 2).

4.1.3. Procedure

Participants were assigned to one of four experimental conditions varying in presentation duration (1 s vs. 5 s) and stimulus type (IAPS vs. paintings). Participants rated on 7-point scales whether the image was a figure-ground (1) or a complex scene (7) composition. Another rating referred to the number of objects shown in the picture ((1) "very few"…"a great many" (7)). One rating referred to the (dis)organization of these elements, and participants were asked to report the difficulty of integrating these objects into a coherent scene, or in the case of a figure-ground representation, the difficulty of integrating the parts of the object into one object ((1) "very easy"…"very difficult" (7)). The order of the three ratings was counterbalanced within and across each condition. The inter-trial interval was set to 5 s, and the ratings were self-paced. Two practice trials were run, and it was made sure that the participants understood the specific terms used in the instructions. Moreover, the word "complexity" never appeared in the instructions to ensure that participants were unaware of the purpose of the study. As in Experiments 1 and 2, the pictures were shown in four blocks and participants could take a break if necessary. After the experiment, participants filled the questionnaire on general aspects of the experiment and one on visual art interest. The experimental sessions lasted approximately 40–60 min, depending on the time condition, and participants were debriefed and dismissed afterwards.

4.2. Results

The results section comprises an analysis of the relationships between the subdimensions of visual complexity with general complexity ratings obtained in Experiments 1 and 2. This analysis is followed by linear regression analyses to predict general complexity ratings by the number of objects and their disorganization. Next, the effect of presentation duration on the averaged ratings of each subdimension is assessed. Last, these subjective ratings are related to the objective measures of complexity and compared across the different conditions of presentation duration.

4.2.1. Change of subjective ratings over time

The four groups were characterized by a similar degree of interest in visual arts, $F(3105) = 1.31$, $p = .276$. Normality checks of the questionnaire data on participants' liking for the stimuli used in the experiment, the level of difficulty of giving judgments as well as the frequency of having seen the paintings prior to the experiment revealed that all except one Kolmogorov–Smirnov normality tests were significant (all $ps < .093$). For IAPS pictures, a series of Mann–Whitney U-tests revealed that the answers did not vary between the two presentation duration conditions (all $ps > .263$), suggesting that presentation duration did not significantly affect the liking for the pictures ($M = 4.18$, $SD = 1.26$) and the degree of experienced emotions ($M = 5.4$, $SD = 1.03$), as well as the difficulty of judging the number of objects ($M = 3.0$, $SD = 1.31$), the integration of these objects into scenes ($M = 4.05$, $SD = 1.9$) and the differentiation between a figure-ground and a complex scene composition ($M = 3.79$, $SD = 1.36$). The familiarity with the IAPS picture set was low ($M = 2.07$, $SD = 1.19$) and similar in both conditions of presentation duration. A similar analysis was conducted for paintings, and only one significant difference on these measures was observed. The difficulty of judging the number of objects in paintings was lower in the 1-s condition ($M = 2.58$, $SD = 1.36$) than in the 5-s condition ($M = 3.46$, $SD = 1.27$), Mann–Whitney $U = 473.0$, $p = .011$, $r = -.35$. The familiarity with the paintings was low ($M = 2.69$, $SD = 1.23$). The paintings ($M = 5.17$, $SD = 1.13$) were liked more than IAPS pictures, but the degree of experienced emotions ($M = 4.65$, $SD = 1.37$) was lower for paintings. The difficulty of judging the integration of objects into scenes ($M = 4.46$, $SD = 1.63$) and the differentiation between a figure-ground composition and a complex scene ($M = 4.02$, $SD = 1.38$) was comparable to the reports regarding IAPS pictures.

Before averaging ratings across each picture, inter-rater reliability was assessed for each condition. In general, inter-rater reliability was high in all conditions. A comparison of all conditions revealed that the lowest ICC values were obtained for the ratings of the difficulty of integrating objects into a coherent scene, or in the case of a figure-ground representation, the difficulty of integrating the parts of the object into one object. The results for IAPS pictures presented for 1 s were as follows: figure-ground vs. complex scene composition (ICC(2, 27) = .98, 95% confidence interval (CI) [.97, .99]), number of objects (ICC(2, 27) = .99, 95% CI [.98, .99]), and disorganization (ICC(2, 27) = .87, 95% CI [.83, .90]). Similar results were obtained for the presentation duration of 5 s: figure-ground vs. complex scene composition (ICC(2, 31) = .98, 95% CI [.98, .99]), number of objects (ICC(2, 31) = .99, 95% CI [.98, .99]), and disorganization (ICC(2, 31) = .83, 95% CI [.78, .88]). Inter-rater reliability was also assessed for paintings in the 1-s condition: figure-ground vs. complex scene composition (ICC(2, 26) = .98, 95% CI [.97, .99]), number of objects (ICC(2, 26) = .98, 95% CI [.98, .99]), and disorganization (ICC(2, 26) = .90, 95% CI [.87, .93]). The results presented for paintings presented for 5 s were as follows: figure-ground vs. complex scene composition (ICC(2, 26) = .98, 95% CI [.98, .99]), number of objects (ICC(2, 26) = .99, 95% CI [.98, .99]), and disorganization (ICC(2, 26) = .80, 95% CI [.74, .86]).

Kolmogorov–Smirnov normality tests indicated that all subjective ratings of IAPS pictures and paintings, except the four on disorganization, significantly deviated from normality. Therefore, it was decided to conduct non-parametric correlations to probe the relationships between general complexity ratings of the images obtained in Experiments 1 and 2 and the ratings of this experiment: figure-ground vs. complex scene composition, number of objects and disorganization. Table 7 shows that for IAPS pictures presented for 1 s, the correlations of these three ratings with general complexity (Experiment 1) were positive and similarly strong with $r_s \sim .70$. For the presentation duration of 5 s, the correlations with figure-ground vs. complex scene composition and number of objects were stronger ($r_s \sim .77$) than for disorganization ($r_s = .51$). For paintings presented for 1 s, the correlations between

Table 7
Spearman's rank order correlations between ratings of general complexity, the differentiation between a figure-ground and complex scene composition, the number of objects and their disorganization in response to IAPS pictures and representational paintings (ART) for two conditions of presentation duration.

Measure	Condition	Complexity	FG-Complex scene	Number of objects
FG-Complex scene	1 s IAPS	.72* (91)		
	5 s IAPS	.77* (94)		
	1 s ART	.63* (89)		
	5 s ART	.75* (90)		
Number of objects	1 s IAPS	.65* (89)	.93* (92)	
	5 s IAPS	.76* (93)	.92* (93)	
	1 s ART	.52* (87)	.92* (91)	
	5 s ART	.69* (90)	.93* (94)	
Disorganization	1 s IAPS	.70* (91)	.69* (94)	.62* (92)
	5 s IAPS	.51* (92)	.63* (92)	.62* (91)
	1 s ART	.54* (87)	.83* (92)	.78* (89)
	5 s ART	.52* (87)	.69* (90)	.70* (90)

Note. $p^* < .05$ after Bonferroni–Holm correction. Degrees of freedom are in parentheses.

general complexity and figure-ground vs. complex scene composition were somewhat stronger ($r_s = .63$) than those for number of objects and disorganization ($r_s \sim .53$). However, in the 5-s condition, the correlations for figure-ground vs. complex scene composition and number of objects were somewhat stronger ($r_s \sim .71$) than for disorganization ($r_s \sim .52$). We observed the same pattern of results as in IAPS pictures presented for 5 s, which may indicate a general trend showing that disorganization is less correlated with general complexity ratings in the 5-s condition. This claim is supported by the fact that the strength of the correlations for IAPS pictures and paintings is similar for this time condition in both picture sets.

Table 7 also offers insights into the inter-relationships between ratings of figure-ground vs. complex scene composition, number of objects and disorganization. These inter-relationships seemed to be less affected by presentation duration. As expected, in both IAPS pictures and paintings the ratings of figure-ground vs. complex scene composition and number of objects were highly correlated in both conditions of presentation duration ($r_s \sim .92$). In IAPS pictures, the correlations between figure-ground vs. complex scene composition and disorganization and the correlations between number of objects and disorganization were moderate (all $r_s \sim .63$) and also similar for both presentation durations. However, the correlations for paintings were stronger in the 1-s condition ($r_s \sim .80$) than in the 5-s condition ($r_s \sim .70$). This may suggest that during the perception of art the inter-correlations between disorganization and figure-ground vs. complex scene composition and the number of objects, respectively, are stronger for shorter presentation durations. This is not the case for environmental scenes, in which the inter-correlations are similar and seem to be unaffected by presentation duration.

Four linear stepwise multiple regression analyses were computed to predict general complexity ratings obtained in Experiments 1 and 2 by ratings of the number of objects and their disorganization. Ratings of figure-ground vs. complex scene were excluded due to the presence of multicollinearity (see Table 8) and on theoretical grounds (e.g., Nadal et al. (2010) did not include this type of rating). Table 8 shows that in three out of four regression models, the number of objects was the sole significant predictor of general complexity ratings. For IAPS pictures in the 1-s condition, both the number of objects and their disorganization were predictors of general complexity. For paintings, the model of the 5-s condition explained more variance (46%) than the model based on the 1-s condition (26%). For IAPS pictures, the results of the models (53% vs. 53%) are less comparable because the number of predictors differs, and disorganization was a stronger predictor in the model of the 1-s condition.

Table 8
Summary of stepwise regression analyses predicting general complexity ratings by the number of objects and their disorganization, separately given for IAPS pictures and representational paintings and two presentation durations.

	Variable	*B*	*SE B*	*β*
IAPS 1 s	Step 1			
($N = 91$)	Constant	2.30	.18	
	Disorganization	.51	.06	.68***
	Adjusted R^2	.45		
	F	75.15***		
	Step 2			
	Constant	2.33	.17	
	Disorganization	.36	.07	.49***
	Objects	.13	.03	.34***
	Adjusted R^2	.53		
	F	51.02***		
	ΔR^2	.08		
IAPS 5 s	Step 1			
($N = 93$)	Constant	3.0	.11	
	Objects	.32	.03	.73***
	Adjusted R^2	.53		
	F	103.56***		
Paintings 1 s	Step 1			
($N = 87$)	Constant	3.45	.12	
	Objects	.20	.04	.52***
	Adjusted R^2	.26		
	F	31.0***		
Paintings 5 s	Step 1			
($N = 89$)	Constant	3.31	.13	
	Objects	.30	.04	.68***
	Adjusted R^2	.46		
	F	74.34***		

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Next, we examined how the three types of complexity ratings changed over time. Fig. 6 shows a comparison of the mean ratings of figure-ground vs. complex scene composition, the number of objects and their disorganization in two conditions of presentation duration, depicted separately for IAPS pictures and representational paintings. With regard to the differentiation between a figure-ground and a complex scene composition, the majority of the IAPS pictures and paintings were rated as looking more like a complex scene in the 5-s condition than in the 1-s condition. Fifty-eight IAPS pictures received higher ratings in the 5-s condition, 37 lower ratings and one was rated the same ($T = 1645$, $p = .018$, $r = -.24$). Similarly, fifty-nine paintings were judged as being more like complex scene, 33 as less like a complex scene, and the rating of 4 pictures did not change ($T = 1310$, $p = .001$, $r = -.33$). Note that paired t-tests yielded the same results.

The data also suggests (Fig. 6) that ratings of the number of elements and their disorganization differ less in the 1-s than in the 5-s condition. In IAPS pictures, the difference between ratings of the number of objects and their disorganization is less pronounced (Wilcoxon signed-rank test, $T = 2124.5$, $p = .815$, $r = -.02$) in the 1-s condition than in the 5-s condition ($T = 1662.5$, $p = .064$, $r = -.19$). In paintings, a similar pattern of results emerged, with a smaller difference between the ratings in the 1-s condition ($T = 1938.5$, $p = .935$, $r = -.08$) than in the 5-s condition ($T = 1687$, $p = 078$, $r = -.18$). To further examine the nature of these larger differences in the 5-s condition in both picture sets, Wilcoxon signed-rank tests indicated that ratings of the number of objects in IAPS pictures varied significantly ($T = 1413.5$, $p = .002$, $r = -.36$), with 67 out of 94 pictures being rated as having fewer objects in the 5-s condition than in the 1-s condition. In a similar vein, 71 out of 94 IAPS pictures were rated as being more organized in the 5-s condition ($T = 536$, $p < .001$, $r = -.66$). However, the results are less clear for paintings. Fifty-seven out of 93 paintings were rated as having more objects in the 5-s condition, 33 were judged as having fewer objects, and for three the ratings did not change ($T = 1252$, $p = .001$, $r = -.33$). For 48 out of 90 paintings ratings of disorganization decreased in the 5-s condition, whereas for 36 they increased, and for 6 they did not change ($T = 1240$, $p = .015$, $r = -.26$).

4.2.2. Changes of the relationships between subjective and objective complexity measures

To analyze the correlations between objective measures of complexity and subjective ratings of figure-ground vs. complex scene composition, the number of objects and their disorganization were computed for IAPS pictures (Table 9) and for representational paintings (Table 10), respectively. The results are discussed by grouping the objective measures according to their type, the picture type and the presentation duration. In the IAPS 1-s condition, correlations between measures of compressed file size and ratings of figure-ground vs. complex scene composition and the number of objects were on average stronger (>~.1) than those with general complexity ratings. The correlations between compressed file sizes and the disorganization ratings were on average weaker (<~.1) than those for general complexity ratings. This pattern of results was also valid for measures of perimeter and Canny edge detection. For the entropy measure, these increases for the correlations with figure-ground vs. complex scene composition and the number of objects were somewhat larger (>~.3), and also the decrease for the correlation with disorganization (<~.2). The correlations for the RMS contrast measures also followed this general pattern,

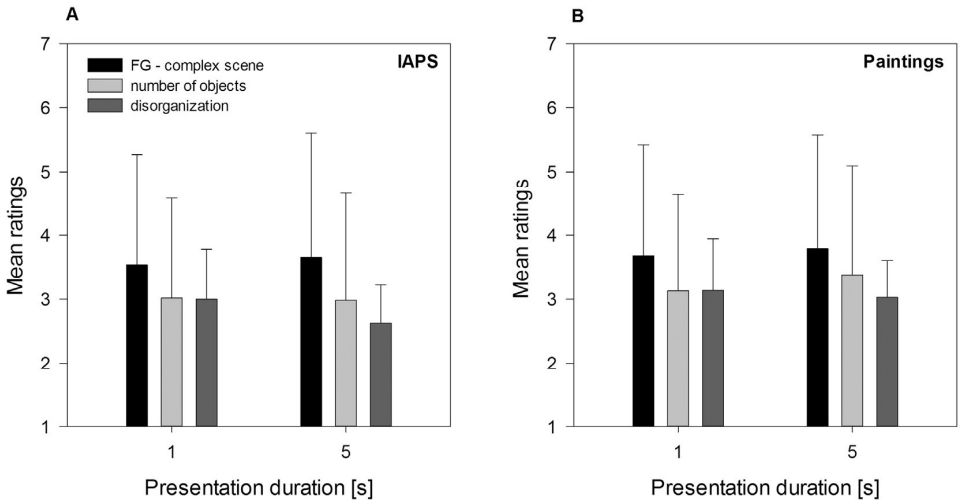

Fig. 6. Mean ratings of figure-ground vs. complex scene composition, number of objects, and disorganization for presentation durations of 1 s and 5 s in response to a) environmental scenes and b) representational paintings. Error bars denote one standard deviation.

but the increases were larger than .1 for the correlations between the mean RMS contrast measure and the ratings of figure-ground vs. complex scene composition and the number of objects (>~.2), respectively. The only measures that deviated from this pattern of results were those based on phase congruency. The correlations for the figure-ground vs. complex scene composition and the number of objects were similar as those observed for general complexity. The decrease in correlation strength for the disorganization ratings was around .2. As expected, the set of objective measures generally correlated stronger with ratings of the number of objects and the differentiation between figure-ground and complex scene composition than with ratings of disorganization. General complexity ratings yielded correlation strengths that lay between the lowest and highest correlations observed for specific types of visual complexity.

In the IAPS 5-s condition (Table 9), the correlations between objective measures and the subdimensions of complexity were of similar strength as those in the 1-s condition. In other words, presentation duration seemed to have less of an effect on correlation strength when subdimensions of complexity were rated in comparison to the condition in which general complexity ratings were given (Experiment 1). In the latter condition, correlation strength slightly increased with longer presentation duration. This finding is also reflected in the fact that increases in correlation strength between objective measures and the subdimensions of complexity observed for the 1-s condition were mostly absent in the 5-s condition. However, the decreases of correlation strength with regard to disorganization ratings were present for most objective measures and particularly strong for the RMS contrast measures. Only for the phase congruency measures, the pattern of results for the 1-s and 5-s conditions was very similar.

For representational paintings (Table 10), the pattern of results was less coherent than the one for IAPS pictures. In comparison to the correlations between general complexity ratings and measures of compressed file size in the 1-s condition, ratings of the number of objects correlated less with compressed file sizes (<~.1), whereas the correlations for disorganization and ratings of figure-ground vs. complex scene composition were similar. The perimeter and Canny edge detection measures as well as the entropy measure yielded correlation strengths with the subdimensions of complexity that were similar to those with general complexity. The mean RMS contrast measure correlated with a similar strength with general complexity and disorganization, but less strongly with the number of objects and ratings of figure-ground vs. complex scene composition (<~.1). However, the *SD* RMS contrast measure correlated the strongest with ratings of disorganization (>~.1) in comparison with the other types of ratings, which also holds true for the alpha RMS contrast measure. Correlations based on the three subdimensions did not lead to stronger correlation strengths for the mean measure of phase congruency, but particularly the *SD* measure of phase congruency correlated well with the number of objects and ratings of figure-ground vs. complex scene composition (>~.1), and to a lesser degree, with disorganization and general complexity. The beta measure led to similar correlations across all types of subjective ratings.

In the 5-s condition, correlations with the measure of JPEG file size slightly improved when ratings of subdimensions of complexity were considered compared to general complexity ratings. The GIF

Table 9
Spearman-Rho correlations between a set of thirteen measures of objective complexity and ratings of general complexity, number of objects, their disorganization and figure ground vs. complex scene in response to IAPS pictures for two presentation durations.

Measure		JPEG	GIF	PNG	TIFF	PERI-RAW	CANNY-RAW	ENTROPY	RMS-CONTR *M*	RMS-CONTR *SD*	RMS-CONTR *α*	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG *β*
Complexity	1 s	.40	.20	.33	.38	.29	.12	.06	.38	.34	.41	.22	.26	.24
(Experiment 1)	5 s	.46	.24	.38	.46	.39	.19	.23	.55	.47	.55	.29	.33	.32
Number of objects	1 s	.52*	.34*	.43*	.49*	.37*	.26	.31	.55*	.48*	.56*	.25	.29	.26
	5 s	.50*	.31	.42*	.49*	.37*	.27	.29	.55*	.48*	.56*	.26	.30	.27
Disorganization	1 s	.28	.22	.22	.28	.19	.16	.23	.27	.15	.23	.03	−.005	.04
	5 s	.30	.27	.23	.26	.22	.18	.27	.30	.16	.25	.11	.10	.11
FG-complex scene	1 s	.51*	.34*	.43*	.50*	.37*	.27	.37*	.56*	.43*	.55*	.23	.25	.24
	5 s	.46*	.29	.38*	.46*	.33*	.27	.37*	.53*	.43*	.53*	.24	.25	.25

Note. $p^* < .05$ after Bonferroni-Holm correction, PERI = perimeter, CONTR = contrast, CONG = congruency, RMS = root mean square, FG = figure-ground.

Table 10
Pearson product–moment correlations between a set of thirteen measures of objective complexity and ratings of general complexity, number of objects, their disorganization and figure ground vs. complex scene in response to representational paintings for two presentation durations.

Measure		JPEG	GIF	PNG	TIFF	LG PERI-RAW	CANNY-RAW	LG ENTROPY REV	RMS-CONTR *M*	RMS-CONTR *SD*	LG RMS-CONTR *α*	PHASE-CONG *M*	PHASE-CONG *SD*	PHASE-CONG *β*
Complexity	1 s	.19	.15	.13	.19	.23	−.23	−.17	.28	.36	.32	.27	.39*	.33
(Experiment 2)	5 s	.06	.13	.04	.08	.21	−.24	−.17	.21	.31	.29	.30	.45*	.37*
Number of objects	1 s	.10	−.04	.02	.05	.18	−.26	−.09	.14	.34*	.26	.31	.53*	.40*
	5 s	.12	−.01	.01	.07	.14	−.27	−.10	.13	.31	.24	.30	.52*	.39*
Disorganization	1 s	.22	.16	.16	.21	.29	−.20	−.12	.32	.42*	.44*	.22	.40*	.30
	5 s	.12	.12	.09	.09	.19	−.04	−.002	.20	.22	.24	.10	.22	.14
FG-complex scene	1 s	.15	.08	.06	.11	.20	−.30	−.11	.19	.32	.29	.27	.47*	.35*
	5 s	.12	.05	.04	.08	.18	−.31	−.12	.18	.32	.30	.27	.47*	.35*

Note. $p^* < .05$ after Bonferroni-Holm correction, PERI = perimeter, CONTR = contrast, CONG = congruency, RMS = root mean square, LG = logarithm, FG = figure-ground.

compressed file size correlated in a similar way for disorganization and general complexity, and not at all with the number of objects and the differentiation between a figure-ground and complex scene composition. PNG and TIFF measures did not correlate with any of the four types of subjective measures. The correlations with the perimeter measure were similar for all types of subjective ratings. The measure based on the Canny algorithm was lower for the correlation with disorganization (<~.2) and of similar strength for the correlations with general complexity, the number of objects and ratings of figure-ground vs. complex scene composition, respectively. A similar pattern was observed for the entropy measure. In general, correlations with the RMS contrast measures did not vary much across the four types of ratings. The three phase congruency measures correlated weaker with disorganization (<~.2) in comparison to the three other types of ratings.

A comparison between the results obtained for IAPS pictures and representational paintings showed that ratings of subdimensions of complexity generally only yielded stronger correlation strengths in IAPS pictures. Only for some measures, such as those based on the Canny algorithm, RMS contrast or phase congruency, increases in correlation strength for certain subdimensions were observed in representational paintings. Furthermore, regarding the factor of presentation duration the data also suggests that in most cases the correlation strength with a particular subdimension of complexity was unaffected by presentation duration. This stands in contrast to results reported in Experiments 1 and 2. In representational paintings, exceptions from this pattern of results refer to measures based on the Canny algorithm, the RMS contrast and phase congruency as well as to entropy, and their relationship to disorganization ratings. In all cases, the correlations were weaker or absent in the 5-s condition than in the 1-s condition. A similar finding was also observed for IAPS pictures and the phase congruency measures, which were not correlated to disorganization ratings in the 1-s compared to the 5-s condition.

In summary, we found indications for an emergence of the multidimensionality of visual complexity in both picture sets by examining the change of the inter-relationships between general complexity ratings and ratings of subdimensions of complexity. Moreover, average ratings of the number of objects and their disorganization only differed in the 5-s but not in the 1-s condition, supporting the findings observed by examining correlation strengths. In three out of four cases, regression analyses showed that the number of objects was the strongest predictor for general complexity ratings. Regarding the relationships between subdimensions of complexity and objective measures, the current findings pointed to a sensitivity of objective measures to the type of subdimension of complexity. These correlations were less prone to be affected by presentation duration than correlations with general complexity ratings, particularly in environmental scenes.

4.3. Discussion

An analysis of the relationships between general complexity ratings and ratings of figure-ground vs. complex scene composition, the number of objects and disorganization has shown that presentation duration affected these correlations in IAPS pictures and representational paintings (Table 7). In the 1-s condition, these correlation strengths for IAPS pictures and representational paintings were similar and not differentiated. However, in the 5-s condition general complexity correlated similarly well with ratings of figure-ground vs. complex scene and the number of objects, respectively. These correlations were higher than the ones with disorganization, regardless of the stimulus type. This finding is also reflected in Fig. 6, where mean ratings of the number of objects and disorganization were similar in the 1-s but not in the 5-s condition. Taken together, this suggests that the multidimensionality of complexity varies over time, especially since the pattern of results was present in both stimulus types.

Furthermore, the experiment demonstrated that the interrelationships between the ratings of subdimensions of complexity were moderately to highly correlated and in all except one case unaffected by presentation duration. In particular, the correlations between the number of objects and the differentiation between a figure-ground composition were highly correlated in IAPS pictures and paintings (all $r_s > .9$), demonstrating that these two types of ratings yielded similar results for the current stimuli sets. This may have theoretical implications because one could argue that the more global assessment along the dimension of figure-ground vs. complex scene, which probably implies the assessment of the size of the background in comparisons to the depicted objects, yields similar results as the assessment of the number of objects, which may involve a sort of counting mechanisms. Future studies could therefore examine and compare the cognitive processes behind these types of complexity ratings.

Regression analyses (excluding the figure-ground vs. complex scene ratings as a predictor) showed that in all except one condition the number of objects was the only predictor of general complexity. In the IAPS 1-s condition, the model yielded disorganization as the strongest predictor, followed by the number of objects. Overall, we could thus replicate the often reported finding that general complexity ratings are mostly related to the number of objects (e.g., Berlyne et al., 1968; Nadal et al., 2010). Further studies employing other stimuli sets and presentation durations may clarify whether this pattern of results can be generalized.

The current data also offered insights into the question of how subjective ratings of figure-ground vs. complex scene composition, the number of objects and their disorganization change between 1 and 5 s of presentation duration (Fig. 6). In both picture sets, mean ratings of figure-ground vs. complex scene were higher in the 5-s condition than in the 1-s condition, suggesting that the level of complexity increased. This finding is in line with Experiments 1 and 2 (but not with Bachmann and Vipper (1983)), which showed that general complexity ratings increased between 1 and 5 s of presentation duration. However, when ratings of the number of objects and disorganization were considered, the pattern changed for IAPS pictures, which received lower ratings for both types of complexity ratings in the 5-s condition, indicating that the pictures were perceived as simpler. In paintings, ratings of the number of objects increased and those of disorganization

decreased. Disorganization ratings were lower in the 5-s condition in both picture sets, suggesting that it was easier to integrate objects into scenes (or parts of objects into an object) under the longer presentation duration. This is in line with Cupchik and Berlyne (1979), as well as Bachmann and Vipper (1983), who found that orderliness increased with exposure time. Furthermore, it is interesting to point out that in our experiment ratings of figure-ground vs. complex scene composition follow the pattern of general complexity ratings better than ratings of the traditional subdimensions of complexity. This is particularly relevant since the instructions avoided any reference to complexity. Altogether, it can be argued that the time course of subjective complexity perception may depend on how complexity is assessed and possibly also on the stimulus type.

Regarding the performance of objective complexity measures, we found that presentation duration generally did not affect the correlations with ratings of figure-ground vs. complex scene, the number of objects and disorganization in IAPS pictures, and to a lesser extent, in representational paintings. However, presentation duration affected the correlations between general complexity and several objective measures particularly in IAPS pictures. Although the reported effects were rather small, it appears as if more concrete ratings referring to subdimensions of complexity may be advantageous if the performance of an objective measure is compared across different stimulus sets and presentation durations.

Our data also shows that objective measures are sensitive to subdimensions of visual complexity. For example, in IAPS pictures mostly selected to vary along the dimension of figure-ground vs. complex scene, ratings of figure-ground vs. complex scene and the number of objects yielded stronger correlations with objective measures in comparison to ratings of disorganization and general complexity. In paintings, sensitivity of objective measures that performed well (RMS contrast and phase congruency measures) to subdimensions of complexity and presentation duration could be observed as well, with some measures better capturing disorganization ratings, and some better those regarding the number of objects and the differentiation between a figure-ground composition and a complex scene. Although some of the differences in correlation strengths were small, other correlation coefficients varied around .2 for the same measure, thus motivating future studies to replicate these findings with other types of stimuli. Only by extending this type of research will we be able to understand why certain objective measures are better able to capture aspects of complexity than others. Particularly in combination with the study of different presentation durations, this may offer a fruitful way of studying visual perception.

5. General discussion

Our current study adds to the growing body of research investigating how emotional and esthetic responses change over time, particularly beyond the initial stages of Gestalt perception. Here, the focus was placed on complexity in the context of two types of visual stimuli: environmental scenes and representational paintings. Their comparison across stimuli sets enabled insights into the nature of esthetic experiences. We investigated response patterns of subjective ratings across three different presentation durations (1 s, 5 s and 25 s). Moreover, we used a set of objective measures of complexity as a tool to study complexity perception. The present study only used evaluations of visual complexity. In this endeavor, the multidimensionality of visual complexity was also acknowledged and interesting results emerged.

First, our results indicated that subjective experience of familiarity, complexity, arousal and pleasantness, which all constitute common concepts in the study of emotion and esthetic experiences, are affected by presentation duration regardless of the stimulus type. This was observed by comparing average ratings of pictures across different presentation durations, by looking at changes of inter-relationships between these variables and mediation effects within the context of Berlyne's psychobiological model (1971), as well as by studying the time course of the multidimensionality of visual complexity. Thus, measures of complexity for a specific presentation duration might not be generalized to different presentation times. Consequently, these results clearly motivate further research aiming at extending these findings with different stimulus categories presented at a wider range of presentation durations. For example, it remains an open question whether the current results can be generalized to abstract stimuli without motivationally relevant information.

Second, our experiments also revealed several similarities and differences between the perception of complexity in environmental scenes and paintings. For example, ratings of general complexity and pleasantness changed in a similar way (inverted U-shape) across the three presentation durations, and interrelationships between complexity and arousal, controlling for familiarity, were increasing with presentation duration in both stimulus sets. However, changes regarding the interrelationships of variables over time were more pronounced regarding representational paintings in general, which may be a characteristic of esthetic experiences. Similarly, the pattern of results regarding mediation effects and their changes over time was not the same in both stimulus sets, suggesting that different underlying cognitive mechanisms may be present. Consequently, our results contribute to the ongoing discussion about the nature of esthetic experiences (Marković, 2012; Vessel & Rubin, 2010; Vessel et al., 2014) by demonstrating that the study of the time course of subjective responses to different stimulus categories is an effective means to tackle this intriguing issue. Further research may include other types of ratings to explore differences between these stimulus categories, such as ratings referring to beauty, preference, fascination and absorption or ratings referring to semantic processing.

Third, we successfully demonstrated that objective complexity measures, mostly based on compressed file size and edge detection algorithms, can enrich the study of the time course of complexity perception. The correlation between objective measures of complexity with subjective ratings of general complexity largely depends on the stimulus type (Chikhman et al., 2012; Marin & Leder, 2013), and this dependency on the stimulus type did not change across different presentation durations. However, examining the performance of these measures across different presentation durations within one stimulus set showed that correlation strengths increased over time in environmental scenes. In representational paintings, however, the pattern of results was less clear, but some objective measures correlated less in the 25-s condition than in conditions of shorter presentation durations, and some measures did not change at all. This perhaps suggests that at longer presentation durations complexity ratings of paintings may be not so strongly based on the visual structure of the image. Higher-order cognitive interferences in the perception of art, such as associations triggered by the semantic content or the semantic content itself, may influence complexity judgements. A recent study by Jakesch and Leder (2015) addressed this issue and showed that semantically ambiguous paintings by Magritte were rated as more complex than matched unambiguous paintings of the same objective complexity. In other words, complexity judgments may not only be based on bottom-up, quantitative aspects of an image, but be related to qualitative aspects such as semantic content. In this study, we did not consider ambiguity since this was not a factor of interest. However, it cannot be ruled out that paintings are generally more ambiguous than environmental scenes and that this contributed to complexity judgements and their changes over time. It is worthwhile considering these aspects in future studies on perceived complexity and comparisons across stimulus types.

Fourth, we showed that the multidimensionality of subjective complexity may emerge over time. For example, ratings of the number of objects and disorganization in environmental scenes and representational paintings were less differentiated in the 1-s than in the 5-s condition, which extends reports by Nadal et al. (2010) who demonstrated

the multidimensionality of visual complexity for a self-paced presentation time. From this it can be concluded that the study of visual complexity requires a careful consideration of the presentation duration of stimuli. We thus suggest studying the time course of the perception of other dimensions of visual complexity, such as symmetry, aspects of color and 3-D appearance, in a wide range of stimulus types, which may finally help clarify our conception of visual complexity. Specific tests of how complexity and its subdimensions are established and change over time might profit from using well-controlled stimuli such as abstract patterns (see Jacobsen and Höfel (2002); Gartus and Leder (2013)).

Finally, we showed that objective measures of complexity are sensitive to subdimensions of visual complexity and that these correlations are less affected by presentation duration than those with general complexity ratings. Nevertheless, our understanding of why this is the case is sparse because at the moment it is unclear why certain measures of objective perform better than others. We are lacking studies using the same set of objective measures across different stimulus sets. Only then we can understand what aspects of visual complexity these measures, often based on complex algorithms, are capturing. Often the aim is to model subjective complexity by developing new objective measures that perform well on one stimulus set (e.g., Cavalcante et al., 2014) rather than to compare the performance of a set of measures across stimulus sets (e.g., Chikhman et al., 2012).

Several limitations of the current study need to be addressed. We restricted our participant samples to females since Marin and Leder (2013) showed that gender effects occurred with regard to the stimuli under investigation. Therefore, future studies should also test male participants. Moreover, it is unclear whether systematic variations in subjective ratings between 5 and 25 s may occur. Consequently, a wider range of presentation durations should be used in future studies. Another limitation concerns the fact that only visual stimuli with semantic content were studied and that findings with regard to abstract patterns or artistic styles may differ. Thus effects on the time course of complexity perception due to semantic content need be disentangled from those that are purely based on the structure of the visual signal. Furthermore, our research is limited to behavioral and computational methods. We are aware that neurophysiological methods are the next step to understand the underlying biological mechanisms of how and why complexity perception is changing over time.

The present study on visual complexity processing nicely complements other current investigations of dynamical aspects of an esthetic episode. For example, studies employing a set of different presentation durations addressed the microgenesis of style and content (Augustin, Defranceschi, Fuchs, Carbon, & Hutzler, 2011; Augustin, Leder, Hutzler, & Carbon, 2008), the perception of pictorial balance (Locher & Nagy, 1996), or the influence of title information on the appreciation of paintings (Leder, Carbon, & Ripsas, 2006). The development of other methods than changing the presentation durations, such as the repeated presentation of visual stimuli (e.g., Carbon & Leder, 2005; Höfel & Jacobsen, 2003; Tinio & Leder, 2009), the recording of viewing time (e.g., Brieber, Nadal, Leder, & Rosenberg, 2014) as well as continuous physiological or behavioral measurements of affective and cognitive processes over certain time intervals (e.g., Fischer, Graupner, Velichkovsky, & Pannasch, 2013; Muth, Raab, & Carbon, 2015; Noguchi & Murota, 2013) further reflect the current trend to study the dynamics of perceptual and cognitive processes.

6. Conclusions

Complexity is a frequently studied aspect of visual perception in general, and more specifically, within the field of empirical esthetics. In three experiments, we investigated the time course of complexity perception induced by affective environmental scenes and representational paintings by combining subjective ratings with objective measures. Importantly, we demonstrated that the study of the time course of complexity perception highly profits from acknowledging the multidimensionality of complexity and the use of objective measures. As such, our findings have theoretical implications and lend support to the view that results of previous studies based on one presentation duration cannot necessarily be generalized across different presentation durations and different aspects of visual complexity. We thus argue that our approach should eventually be extended to other types of visual stimuli and to a larger set of presentation durations.

Acknowledgments

The authors would like to thank Katrin Mair, Lars-Michael Schöpper, and Denise Ramsperger for their assistance in data collection and Bruno Gingras for his proofreading and critical comments. This research was partly supported by a grant from the Austrian Science Fund (FWF) to HL, project number: P23538-G21. We would like to state that results of Experiments 1 and 2 were partly based on data already published earlier (Marin & Leder, 2013).

References

Aksentijevic, A., & Gibson, K. (2012). Complexity equals change. *Cognitive Systems Research*, *15*, 1–16. http://dx.doi.org/10.1016/j.cogsys.2011.01.002.

American Psychological Association (2010). *Ethical principles of psychologists and code of conduct Retrieved from http://apa.org/ethics/code/index.aspx*

Augustin, M.D., Leder, H., Hutzler, F., & Carbon, C.C. (2008). Style follows content: On the microgenesis of art perception. *Acta Psychologica*, *128*(1), 127–138. http://dx.doi.org/10.1016/j.actpsy.2007.11.006.

Augustin, M.D., Defranceschi, B., Fuchs, H.K., Carbon, C.C., & Hutzler, F. (2011). The neural time course of art perception: An ERP study on the processing of style versus content in art. *Neuropsychologia*, *49*(7), 2071–2081. http://dx.doi.org/10.1016/j.neuropsychologia.2011.03.038.

Bachmann, T., & Vipper, K. (1983). Perceptual rating of paintings from different artistic styles as a function of semantic differential scales and exposure time. *Archiv für Psychologie*, *135*(2), 149–161.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., ... Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. http://dx.doi.org/10.3758/BF03196494.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. US, McGraw-Hill Book Company: McGraw-Hill series in psychology. New York pp. 350. http://dx.doi.org/10.1037/11164-000.

Berlyne, D.E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Berlyne, D.E. (1974). Novelty, complexity, and interestingness. In D.E. Berlyne (Ed.), *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic comparison* (pp. 175–180). Washington: Hemisphere.

Berlyne, D.E., Ogilvie, J.C., & Parham, L.C.C. (1968). The dimensionality of visual complexity, interestingness, and pleasingness. *Canadian Journal of Psychology*, *22*, 376–387. http://dx.doi.org/10.1037/h0082777.

Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Blackwell Publishing Inc, Psychophysiology*, *44*(3), 364–373.

Bookstein, F.L. (1997). *Morphometric tools for landmark data: Geometry and biology*. Cambridge University Press.

Borgomaneri, S., Gazzola, V., & Avenanti, A. (2014). Temporal dynamics of motor cortex excitability during perception of natural emotional scenes. *Social Cognitive and Affective Neuroscience*, *9*(10), 1451–1457. http://dx.doi.org/10.1093/scan/nst139.

Bradley, M.M., Codispoti, M., Sabatinelli, D., & Lang, P.J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, *1*, 300–319.

Bradley, M.M., Houbova, P., Miccoli, L., Costa, V.C., & Lang, P.J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, *48*, 1544–1553. http://dx.doi.org/10.1111/j.1469-8986.2011.01223.x.

Braun, J., Amirshahi, S.A., Denzler, J., & Redies, C. (2013). Statistical image properties of print advertisements, visual artworks and images of architecture. *Frontiers in Psychology*, *4*. http://dx.doi.org/10.3389/fpsyg.2013.00808.

Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PloS One*, *9*(6), e99019. http://dx.doi.org/10.1371/journal.pone.0099019.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-8*(6), 679–698. http://dx.doi.org/10.1109/TPAMI.1986.4767851.

Carbon, C.C. (2012). Dynamics of aesthetic appreciation. *Human Vision and Electronic Imaging*, *8291*(1A), 1–6. http://dx.doi.org/10.1117/12.916468.

Carbon, C.C., & Leder, H. (2005). The Repeated Evaluation Technique (RET). A method to measure dynamic effects of innovativeness and attractiveness. *Applied Cognitive Psychology*, *19*(5), 587–601. http://dx.doi.org/10.1002/acp.1098.

Cavalcante, A., Mansouri, A., Kacha, L., Barros, A.K., Takeuchi, Y., Matsumoto, N., & Ohnishi, N. (2014). Measuring streetscape complexity based on the statistics of local contrast and spatial frequency. *PloS One*, *9*(2), e87097. http://dx.doi.org/10.1371/journal.pone.0087097.s001.

Chatterjee, A. (2003). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of Psychology and the Arts*, *4*(2), 55–60.

Chatterjee, A. (2014). Scientific aesthetics: Three steps forward. *British Journal of Psychology*, *105*(4), 465–467. http://dx.doi.org/10.1111/bjop.12086.

Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, *41*, 631–647. http://dx.doi.org/10.1068/p6987.

Chipman, S.F. (1977). Complexity and structure in visual patterns. *Journal of Experimental Psychology: General*, *106*(3), 269–301.

Codispoti, M., Mazzetti, M., & Bradley, M.M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*, 731–738. http://dx.doi.org/10.1111/j.1469-8986.2009.00804.x.

Cupchik, G.C., & Berlyne, D.E. (1979). The perception of collative properties in visual stimuli. *Scandinavian Journal of Psychology*, *20*(1), 93–104. http://dx.doi.org/10.1111/j.1467-9450.1979.tb00688.x.

Cupchik, G.C., & Gebotys, R. (1988). The experience of time, pleasure, and interest during aesthetic episodes. *Empirical Studies of the Arts*, *6*(1), 1–12. http://dx.doi.org/10.2190/5YN3-J3P8-FWHY-UDB3.

Day, H. (1967). Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. *Perception & Psychophysics*, *2*, 281–286. http://dx.doi.org/10.3758/BF03211042.

Donderi, D.C. (2006a). An information theory analysis of visual complexity and dissimilarity. *Perception*, *35*, 823–835. http://dx.doi.org/10.1068/p5249.

Donderi, D.C. (2006b). Visual complexity: A review. *Psychological Bulletin*, *132*, 73–97.

Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of western and African folk melodies by western and African listeners. *Psychology of Music*, *34*(3), 337–371. http://dx.doi.org/10.1177/0305735606064842.

Eisenman, R. (1968). Semantic differential ratings of polygons varying in complexity–simplicity and symmetry–asymmetry. *Perceptual and Motor Skills*, *26*(3c), 1243–1248. http://dx.doi.org/10.2466/pms.1968.26.3c.1243.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1), 1–29. http://dx.doi.org/10.1167/7.1.10.

Fischer, T., Graupner, S.T., Velichkovsky, B.M., & Pannasch, S. (2013). Attentional dynamics during free picture viewing: Evidence from oculomotor behavior and electrocortical activity. *Frontiers in Systems Neuroscience*, *7*. http://dx.doi.org/10.3389/fnsys.2013.00017.

Forster, M., Leder, H., & Ansorge, U. (2013). It felt fluent, and I liked it: Subjective feeling of fluency rather than objective fluency determines liking. *Emotion*, *13*(2), 280–289. http://dx.doi.org/10.1037/a0030115.

Forsythe, A. (2009). Visual complexity: Is there all there is? In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics. 5639*. (pp. 158–166). Berlin: Springer.

Forsythe, A., Sheehy, N., & Sawey, M. (2003). Measuring icon complexity: An automated analysis. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 334–342. http://dx.doi.org/10.3758/BF03202562.

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic level pictures processing. *Behavior Research Methods*, *40*(1), 116–129. http://dx.doi.org/10.3758/BRM.40.1.116.

Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C.J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, *102*, 49–70. http://dx.doi.org/10.1348/000712610X498958.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*(special issue 2001–2002), 123–147. http://dx.doi.org/10.1177/10298649020050S105.

Gartus, A., & Leder, H. (2013). The small step toward asymmetry: Aesthetic judgment of broken symmetries. *i-Perception*, *4*(5), 361–364. http://dx.doi.org/10.1068/i0588sas.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. http://dx.doi.org/10.1016/S0010-0277(98)00034-1.

Golland, Y., Keissar, K., & Levit-Binnun, N. (2014). Studying the dynamics of autonomic activity during emotional experience. *Psychophysiology*, *51*(11), 1101–1111. http://dx.doi.org/10.1111/psyp.12261.

Gonzalez, R.C., Woods, R.E., & Eddins, S.L. (2003). *Digital image processing using MATLAB*. New Jersey: Prentice Hall.

Graham, D.J., & Field, D.J. (2007). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision*, *21*(1), 149–164. http://dx.doi.org/10.1163/156856807782753877.

Graham, D.J., & Meng, M. (2011). Artistic representations: Clues to efficient coding in human vision. *Visual Neuroscience*, *28*(04), 371–379. http://dx.doi.org/10.1017/S0952523811000162.

Greene, M.R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*, *58*(2), 137–176. http://dx.doi.org/10.1016/j.cogpsych.2008.06.001.

Hayes, A.F., & Preacher, K.J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research*, *45*(4), 627–660. http://dx.doi.org/10.1080/00273171.2010.498290.

Heaps, C., & Handel, C.H. (1999). Similarity and features of natural textures. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 299–320.

Hekkert, P., & van Wieringen, P.C.W. (1990). Complexity and prototypicality as determinants of the appraisal of Cubist paintings. *British Journal of Psychology*, *81*, 483–495. http://dx.doi.org/10.1111/j.2044-8295.1990.tb02374.x.

Höfel, L., & Jacobsen, T. (2003). Temporal stability and consistency of aesthetic judgments of beauty of formal graphic patterns. *Perceptual and Motor Skills*, *96*(1), 30–32. http://dx.doi.org/10.2466/pms.2003.96.1.30.

Ichikawa, S. (1985). Quantitative and structural factors in the judgment of pattern complexity. *Perception & Psychophysics*, *38*(2), 101–109. http://dx.doi.org/10.3758/BF03198846.

Jacobsen, T. (2014). Domain specificity and mental chronometry in empirical aesthetics. *British Journal of Psychology*, *105*(4), 471–473. http://dx.doi.org/10.1111/bjop.12094.

Jacobsen, T., & Höfel, L. (2002). Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual and Motor Skills*, *95*, 755–766. http://dx.doi.org/10.2466/pms.2002.95.3.755.

Jakesch, M., & Leder, H. (2015). The qualitative side of complexity: Testing effects of ambiguity on complexity judgements. *Psychology of Aesthetics; Creativity and the Arts*, *9*(3), 200–205.

Kaplan, S., & Kaplan, R. (1982). *Cognition and environment: Functioning in an uncertain world*. New York: Praeger.

Kendall, D.G. (1984). Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, *16*(2), 81–121. http://dx.doi.org/10.1112/blms/16.2.81.

Kovesi, P. (1999). Image features from phase congruency. *Videre: A Journal of Computer Vision Research*, *1*(3), 2–26.

Kovesi, P. (2003). Phase congruency detects corners and edges. In C. Sun, H. Taöbot, S. Ourselin, & T. Adriaansen (Eds.), *Proceedings of the VIIth Biennial Australian Pattern Recognition Society Conference: DICTA 2003* (pp. 309–318). Csiro: Collingwood.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report a-7*. Gainesville, FL: University of Florida.

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode–Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443–464. http://dx.doi.org/10.1111/bjop.12084.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, *95*, 489–508. http://dx.doi.org/10.1348/0007126042369811.

Leder, H., Carbon, C.C., & Ripsas, A.L. (2006). Entitling art: Influence of title information on understanding and appreciation of paintings. *Acta Psychologica*, *121*(2), 176–198. http://dx.doi.org/10.1016/j.actpsy.2005.08.005.

Leder, H., Gerger, G., Brieber, D., & Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cognition & Emotion*, *28*(6), 1137–1147. http://dx.doi.org/10.1080/02699931.2013.870132.

Lithari, C., Frantzidis, C.A., Papadelis, C., Vivas, A.B., Klados, M.A., Kourtidou-Papadeli, C., ... Bamidis, P.D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, *23*(1), 27–40. http://dx.doi.org/10.1007/s10548-009-0130-5.

Locher, P., & Nagy, Y. (1996). Vision spontaneously establishes the percept of pictorial balance. *Empirical Studies of the Arts*, *14*(1), 17–31. http://dx.doi.org/10.2190/X8U3-CTQ6-A7J1-8JQ8.

Locher, P., Overbeeke, K., & Wensveen, S. (2010). Aesthetic interaction: A framework. *Design Issues*, *26*, 70–79. http://dx.doi.org/10.1162/DESI_a_00017.

Marin, M.M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PloS One*, *8*(8), e72412. http://dx.doi.org/10.1371/journal.pone.0072412.

Marković, S. (2012). Components of aesthetic experience: Aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *iPerception*, *3*(Special issue - Art and Perception), 1–17. http://dx.doi.org/10.1068/i0450aap.

Marković, S. (2014). Object domains and the experience of beauty. *Art & Perception*, *2*(1–2), 119–140. http://dx.doi.org/10.1163/22134913-00002020.

Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological model. *American Journal of Psychology*, *103*, 53–80. http://dx.doi.org/10.2307/1423259.

Martini, N., Menicucci, D., Sebastiani, L., Bedini, R., Pingitore, A., Vanella, N., ... Gemignani, A. (2012). The dynamics of EEG gamma responses to unpleasant visual stimuli: From local activity to functional connectivity. *NeuroImage*, *60*(2), 922–932. http://dx.doi.org/10.1016/j.neuroimage.2012.01.060.

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30.

Monin, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology*, *85*(6), 1035–1048.

Muth, C., Raab, M.H., & Carbon, C.C. (2015). The stream of experience when watching artistic movies. Dynamic aesthetic effects revealed by the Continuous Evaluation Procedure (CEP). *Frontiers in Psychology*, *6*(365). http://dx.doi.org/10.3389/fpsyg.2015.00365.

Nadal, M., Munar, E., Marty, G., & Cela-Conde, C.J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts*, *28*(2), 173–191. http://dx.doi.org/10.2190/EM.28.2.d.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. http://dx.doi.org/10.1016/0010-0285(77)90012-3.

Noguchi, Y., & Murota, M. (2013). Temporal dynamics of neural activity in an integration of visual and contextual information in an esthetic preference task. *Neuropsychologia*, *51*(6), 1077–1084. http://dx.doi.org/10.1016/j.neuropsychologia.2013.03.003.

Palumbo, L., Odgen, R., Makin, A.J., & Bertamini, M. (2014). Examining visual complexity and its effect on perceived duration. *Journal of Vision*, *14*, 1–18. http://dx.doi.org/10.1167/14.14.3.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A*, *7*(10), 2032–2040. http://dx.doi.org/10.1364/JOSAA.7.002032.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3.

Redies, C., Amirshahi, S.A., Koch, M., & Denzler, J. (2012, January). PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects.

Computer Vision–ECCV 2012. Workshops and Demonstrations (pp. 522–531). Berlin Heidelberg: Springer.
Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
Rozin, P., & Royzman, E.B. (2001). Negativity bias, negative dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2.
Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.
Schooler, C. (1984). Psychological effects of complex environments during the life span. A review and theory. *Intelligence*, *8*(4), 259–281. http://dx.doi.org/10.1016/0160-2896(84)90011-4.
Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420.
Silvia, P.J. (2009). Looking past pleasure: Anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 48–51. http://dx.doi.org/10.1037/a0014632.
Silvia, P.J., & Brown, E.M. (2007). Anger, disgust, and the negative aesthetic emotions: Expanding on appraisal model of aesthetic experience. *Psychology of Aesthetics, Creativity, and the Arts*, *1*(2), 100–106. http://dx.doi.org/10.1037/1931-3896.1.2.100.
Smith, J.K., & Smith, L.F. (2001). Spending time on art. *Empirical Studies of the Arts*, *19*(2), 229–236. http://dx.doi.org/10.2190/5MQM-59JH-X21R-JN5J.
Smith, L.F., Bousquet, S.G., Chang, G., & Smith, J.K. (2006). Effects of time and information on perception of art. *Empirical Studies of the Arts*, *24*(2), 229–242. http://dx.doi.org/10.2190/DJM0-QBDW-03V7-BLRM.
Terwilliger, R.F. (1963). Pattern complexity and affective arousal. *Perceptual and Motor Skills*, *17*(2), 387–395. http://dx.doi.org/10.2466/pms.1963.17.2.387.
Tinio, P.P. (2013). From artistic creation to aesthetic reception: The mirror model of art. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 265–275. http://dx.doi.org/10.1037/a0030872.
Tinio, P.P., & Leder, H. (2009). Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica*, *130*(3), 241–250. http://dx.doi.org/10.1016/j.actpsy.2009.01.001.
Tinio, P.P., Gerger, G., & Leder, H. (2013). Birds of a feather… Generalization of facial structures following massive familiarization. *Acta Psychologica*, *144*(3), 463–471. http://dx.doi.org/10.1016/j.actpsy.2013.08.003.
Tuch, A.N., Bargas-Avila, J.A., Opwis, K., & Wilhelm, F.H. (2009). Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies*, *67*(9), 703–715. http://dx.doi.org/10.1016/j.ijhcs.2009.04.002.
Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social emotional development. *Psychological Bulletin*, *134*(3), 383–403. http://dx.doi.org/10.1037/0033-2909.134.3.383.
Vessel, E.A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *Journal of Vision*, *10*(2), 1–14. http://dx.doi.org/10.1167/10.2.18.
Vessel, E.A., Stahl, J., Maurer, N., Denker, A., & Starr, G.G. (2014). Personalized visual aesthetics. *Proc. SPIE* 9014. *Human Vision and Electronic Imaging XIX*, *90140S*, 1–8. http://dx.doi.org/10.1117/12.2043126.
Vitz, P.C. (1966). Preference for different amounts of visual complexity. *Behavioral Science*, *11*(2), 105–114. http://dx.doi.org/10.1002/bs.3830110204.
Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch, & K.C. Klauer (Eds.), *The psychology of evaluation: affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum.
Wundt, W.M. (1874). *Grundzüge der physiologischen Psychologie*. Leipzig: Engelmann.

2.3.4 Summary and outlook

In Experiment 1, environmental scenes were presented for either 1, 5 or 25 s to 60 females who reported familiarity, complexity, pleasantness and arousal in a between-subjects design. The distribution of the scenes in the emotion space spanned by pleasantness and arousal varied only mildly over the three presentation durations. The pleasantness ratings of the respective presentation durations correlated more strongly with each other than arousal ratings. The positive correlation between familiarity and pleasantness was similar in strength (r_s = .70) for all presentation durations, whereas the negative correlation between familiarity and complexity (r_s = -.30) was similar for the presentation durations of 1 and 25 s, but not present in the 5-s condition. The longer the presentation duration, the weaker got the negative relationship between familiarity and arousal (r_s = -.52 to r_s = -.35). When controlling for familiarity, the positive correlation between complexity and arousal was only significant in the 5- and 25-s conditions (r_s = .27 and r_s = .21). The positive correlation between complexity and pleasantness was only significant in the 5s-condition.

Having a look at the mean ratings over all pictures revealed an increase in familiarity over time, and arousal ratings were higher in the 5- and 25-s conditions than in the 1-s condition. Pleasantness ratings were the highest in the 5-s condition and lower in the 1- and 25-s conditions (not in line with Codispoti et al., 2009). Similarly, complexity ratings also followed an inverted U-shaped relationship across the three presentation durations, and the lowest ratings were observed for the 25-s condition. Mediated regression analyses revealed that only in the 5- and 25-s conditions arousal was a mediator in the complexity-pleasantness relationship. An increase in complexity was associated with a decrease in pleasantness through an effect of complexity on arousal (R^2 between .53 and .64). Note that complexity and arousal did not significantly correlate with each other in the 1-s condition. Correlations between objective measures and subjective complexity showed that the performance of the respective measures was similar to those reported by Marin and Leder (2013), and further, that presentation duration had no significant effect on these relationships. However, we observed a non-significant increase in the strength of these relationships with increasing presentation durations, which may be due to more “accurate” subjective complexity ratings in the 5-s condition compared to the 1-s condition.

In Experiment 2, 60 other female participants rated representational paintings presented for either 1, 5 or 25 s using the same procedure as in Experiment 1. Re-

garding the emotion space covered by the representational paintings one can conclude that the quadratic relationship between pleasantness and arousal was weaker for paintings than for environmental scenes, and that as in Experiment 1, the correlations between pleasantness ratings across the three presentation durations were stronger than those for arousal. The three emotion spaces covered by the paintings differed to a larger extent from each other than the three emotion spaces for environmental scenes. The shapes of the emotion space of the 1-s condition and the 25-s condition were more similar to each other than the spaces of the 1-s and 5-s conditions. However, the highest similarity in shape was observed for the 5-s and 25-s conditions. A look at the inter-relationships between the four subjective measures revealed that the strength of the relationship between two variables was generally time-dependent, with stronger correlations for longer presentation durations (except for the familiarity-arousal relationship). Notably, this pattern of results was not present for environmental scenes. When controlling for familiarity, the negative complexity-pleasantness relationship did not reach significance in any of the three time conditions. Next, the results indicated that mean familiarity ratings differed over time, with higher ratings in the 1- and 25-s conditions than in the 5-s condition. Complexity ratings were the highest in the 5-s condition with lower ratings for the 1-s and 25-s conditions. Arousal ratings were higher in the 1-s and 25-s conditions than in the 5-s condition. And finally, pleasantness ratings were the highest in the 5-s conditions and lower in the two other conditions. Mediated regression analyses, to test whether arousal was a mediator in the complexity-pleasantness relationship, yielded the following results (R^2 between .57 and .73): complexity and arousal were correlated in all conditions of presentation duration. Indirect effects were observed for paintings of moderate and high complexity for all presentation durations, but not for simple paintings. Regarding the relationship between subjective complexity and a set of objective measures, we were able to replicate the findings of Marin and Leder (2013). Compression rate was not a significant predictor of subjective complexity (because a uniform background can be complex due to the pattern generated by the brush strokes), but edge detection measures, such as the Canny algorithm, RMS contrast and phase congruency measures, again performed better. Presentation duration had no significant effect on these findings; however, some measures (perimeter, entropy, RMS contrast) seemed to perform somewhat better under shorter presentation durations. As observed in Marin and Leder (2013), arousal also correlated with some measures

of objective complexity, with some indications that presentation duration may affect these relationships.

When comparing results of Experiment 1 with those of Experiment 2, one interesting finding is that the inter-relationships between familiarity, complexity, pleasantness and arousal for ratings of paintings increase in strength when the presentation duration is increased. This may be indicative of the dynamics of aesthetic experiences (Leder et al., 2004; Pelowski et al., 2017), in which cognitive and affective factors interact with each other, and it seems that different exposure times and the study of the inter-relationships of core variables may offer a window into the emergence of aesthetic experiences (Brieber et al., 2020; Sano et al., 2025; Schwabe et al., 2018; Verhavert et al., 2018). In a similar vein, this argument is supported by specific results of the mediated regression analyses, and also by several correlations between subjective complexity and several objective complexity measures, which appear to be stronger for shorter presentation durations. Thus, using the same materials, the study of presentation durations shorter than 1 s and between 5 and 25 s may be a fruitful path for the future (but see Brielmann et al., 2017, for counter-evidence when examining the relationship between pleasure and beauty, and Forster et al., 2016, and Jakesch et al., 2013, who did not report positive effects of stimulus durations shorter than 500 ms on liking when studying fluency effects).

One common finding of Experiments 1 and 2 was that the two-dimensional emotion space of environmental scenes and representational paintings, representing an inverted U-shaped curve, was present for all presentation durations. The quadratic relationship was weaker for representational paintings and the three emotion spaces (associated with each presentation duration) were more dissimilar to each other than what was observed for environmental scenes: an analysis of the shape of the distribution of the stimuli in the two-dimensional pleasantness-arousal space showed that presentation duration had little impact on the distribution of environmental scenes, but the emotion spaces of representational paintings varied more and the emotion spaces of the 5-s and 25-s conditions were the most similar. Whereas for environmental scenes and paintings pleasantness ratings were highly correlated with each other across presentation durations (r ~ .90), arousal ratings were less correlated with each other (environmental scenes r_s ~ .80; paintings r ~ .70). A larger variability for preferences for abstract images compared to environmental scenes has been reported by Vessel and Rubin (2010), and it could be that responses to environmental

scenes are less prone to individual differences due to semantic associations that are likely to be similar across participants. Art perception, even though semantic contents is present, may leave more room to individual thoughts and feelings, which may also be reflected in the reported weaker association between arousal ratings across the three presentation durations in paintings than environmental scenes. Interestingly, a comparison of the mean ratings across the three presentation durations showed that complexity and pleasantness changed in a similar way in the two experiments, namely with lower ratings in the 1- and 25-s conditions compared to the 5-s condition, which yielded an inverted U-shaped pattern. Whether this pattern is typical of affective scenes with semantic contents needs to be determined in future experiments (see also Brieber et al., 2020). It can be hypothesized that long presentation durations may lead to boredom and lower pleasantness ratings, especially for pictures of low complexity. This hypothesis gains support by recent findings by Pelowski et al. (2020) who demonstrated that top-down modulations for longer presentation times can decrease liking ratings of paintings of John Constable (a representative of representational art), particularly in lay viewers with high art interest. Regarding arousal ratings of representational paintings, the observed decrease between the 1- and 5-s conditions was previously reported by Cupchik and Berlyne (1979) when investigating paintings and artificial patterns.

Experiment 3 sought to explore complexity in greater depth and considered two underlying complexity dimensions previously reported in the literature, namely the number of elements/objects and dis(organization) (Nadal et al., 2010), which seemed to be relevant for the current type of visual scenes. In addition, we obtained ratings of how the scenes varied according to figure-ground organization vs. complex scene, which was the main criterion for the selection procedure in Marin and Leder (2013). We were interested in whether the multidimensionality of complexity is present already at the shortest presentation durations of 1 s. Moreover, we wanted to study the association between subjective complexity and its sub-dimensions with objective measures across two presentation durations. One-hundred nine female participants rated either environmental scenes or representational paintings according to different complexity measures in one of the two presentation durations (1 or 5 s). There were no group differences regarding age and art interest.

This experiment offered a number of interesting insights: the correlation between figure-ground vs. complex scene ratings and the number of objects was very

high (r_s >.90) for both stimulus sets and presentation durations, and both of these measures correlated strongly (mostly between r_s ~.60 and r_s ~.75) with general complexity ratings obtained in Experiments 1 and 2. The strength of these relationships was somewhat stronger in the 5-s than in the 1-s condition in both stimulus sets. Regression analyses demonstrated that the number of objects was the main determinant of general complexity (Berlyne et al., 1968; Nadal et al., 2010; Van Geert & Wagemans, 2020), except for environmental scenes in the 1-s condition. This strongly suggests that Marin and Leder (2013) manipulated complexity mostly regarding the number of elements/objects in a visual scene. (Dis)organization correlated also with general complexity ratings, but to a lesser extent (except for the 1-s environmental scene condition, all r_s ~.50). In general, this pattern of results held true for both stimulus sets, but a closer look revealed that general complexity ratings in the 1-s art condition showed the weakest correlations with subdimensions of complexity. This may suggest that during the perception of representational art visual complexity is less clearly differentiated in the 1s-condition, probably due to the fact that participants reported more difficulty when judging the number of objects in the 1-s condition compared to the 5-s condition. This interpretation is supported by another finding, namely that for paintings the inter-correlations between (dis)organization, figure-ground vs. complex scene and the number of objects are stronger in the 1-s than in the 5-s condition. Future research on the dynamics of complexity during art perception will have to clarify whether the emergence of the multidimensionality of visual complexity is specific for representational art by comparing different artistic styles and a wider range of shorter presentation durations.

Another interesting insight we gained from this experiment regarding our set of objective measures of complexity was that ratings referring to subdimensions of visual complexity, such as the number of objects in a scene and the differentiation between figure-ground vs. complex scene, correlated more strongly with objective complexity measures of environmental scenes than general complexity ratings. These correlations concerned the four compression rate measures, perimeter detection and the RMS-contrast measures, and they were basically unaffected by presentation duration. However, general ratings of complexity and their relation to objective measures do somewhat vary across presentation durations (Experiment 1), which suggests that the development and application of objective complexity measures should focus on subdimensions rather than on general complexity ratings. Most current stud-

ies are still trying to model general, global ratings of visual complexity (Carballal et al., 2020; Fernandez-Lozano et al., 2019; Machado et al., 2015; Redies et al., 2012) and ignore its specific subdimensions, but Gartus and Leder (2017) considered the two major factors of visual complexity (i.e., a quantitative and structural factor) in their computational models of visual complexity. For representational paintings, we did not observe many significant correlations after Bonferroni-Holm corrections. Measures based on the Canny algorithm, RMS contrast and phase congruency showed some sensitivity to subdimensions of visual complexity, with several indications of effects of presentation duration (correlations with (dis)organization were weaker in the 5-s than in the 1-s condition).

This study has replicated and extended findings reported by Marin and Leder (2013). By systematically comparing environmental scenes with representational paintings, these experiments demonstrated that studying the effect of presentation duration on the subjective experience of variables such as familiarity, complexity, subdimensions of complexity, pleasantness and arousal can reveal insights into the nature of aesthetic experiences. Our findings are currently not generalizable beyond the female study sample and beyond the types of visual stimuli. Therefore, for instance, stimulus categories without semantic contents should be studied in greater depth to avoid top-down influences on complexity judgements (Jakesch & Leder, 2015). Another avenue for future research is to include color, 3D appearance and symmetry as additional dimensions of complexity (Damiano, et al., 2023; Van Geert & Wagemans, 2020). Moreover, a wider range of presentation durations should be studied to better capture the dynamics of aesthetic experiences. Last, one could also conduct a similar study with music. Studies have shown that brief musical excerpts of 3 s contain already much information relevant for a wide range of affective and cognitive judgements (Plazak & Huron, 2011), and emotional responses to music can be reported to much shorter excerpts of less than 1 s (Bigand, et al., 2005).

2.4 Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music (Marin, Lampatz, Wandl, & Leder, 2016)

2.4.1 Introduction

In the last decades, mixed results have been reported regarding Berlyne's proposed inverted-U relationship between collative variables and hedonic tone. For instance, Nadal et al. (2010) showed that the relationship between visual complexity and beauty depends on the subdimension of complexity. Güçlütürk et al. (2016) reported that individual differences in preferences may explain discrepancies regarding the inverted U-shaped relationship between complexity and liking. Here, we surmised that the relationship between a subdimension of complexity (when held constant) and hedonic tone may depend on the specific measure of *hedonic tone*. Berlyne (1971) subsumed concepts such as beauty, liking and pleasantness under the heading of hedonic tone. However, it is possible that these concepts, although related with each other, show different relationships with collative variables, such as complexity. This hypothesis is supported by neuroscientific findings suggesting that different brain areas are active when processing beauty, liking and pleasantness (e.g., Chatterjee & Vartanian, 2014). Therefore, models of aesthetic experiences based on neuroscientific findings in music (Brattico et al., 2013) and the visual arts (Pelowski, Markey et al., 2017) differentiate between concepts such as beauty, liking/preference and pleasantness/pleasure among other things.

In two experiments, we employed between-subjects designs and asked participants to rate either beauty, pleasantness or liking (alongside arousal, complexity and familiarity) of either environmental scenes, environmental scenes turned into cartoons or representational paintings presented for 5 s. In comparison to Marin and Leder (2013, 2016), we introduced a new category of stimuli (i.e., cartoon-like environmental scenes) that was based on the same affective and semantic contents as environmental scenes but gave the impression of being artistic. This allowed for a better manipulation of the factor aesthetic quality. In Experiment 2, we used excerpts of solo piano music (Experiment 3, Marin & Leder, 2013) cut to 15 s. In both experiments, we better controlled for participant characteristics than in our previous work by measuring mood, art interest and musical sophistication. As in Marin and Leder (2016), we restricted our samples to female participants. We also collected several standard questionnaires on individual differences, which were analyzed in Marin and Leder (2018).

2.4.2 Authorship contribution statement

MMM designed the experiments and created the stimulus library. AL, MW, and MMM collected the data. MMM analyzed and interpreted the data. MMM drafted the work and revised it after feedback from AL, MW, and HL.

Note that ratings for environmental scenes and representational paintings in Experiment 1 were collected as part of Allegra Lampatz' master thesis project, which was co-supervised by MMM and HL. The subjective ratings of cartoons in Experiment 1 were collected by MMM. Experiment 2 was part of the master thesis authored by Michaela Wandl and co-supervised by MMM and HL.

2.4.3 Re-print of the published research article

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 04 November 2016
doi: 10.3389/fnhum.2016.00536

Berlyne Revisited: Evidence for the Multifaceted Nature of Hedonic Tone in the Appreciation of Paintings and Music

Manuela M. Marin[1,2], Allegra Lampatz[1], Michaela Wandl[1] and Helmut Leder[1]*

[1] Department of Basic Psychological Research and Research Methods, University of Vienna, Vienna, Austria, [2] Department of Psychology, University of Innsbruck, Innsbruck, Austria

In his seminal book on esthetics, Berlyne (1971) posited an inverted-U relationship between complexity and hedonic tone in arts appreciation, however, converging evidence for his theory is still missing. The disregard of the multidimensionality of complexity may explain some of the divergent results. Here, we argue that definitions of hedonic tone are manifold and systematically examined whether the nature of the relationship between complexity and hedonic tone is determined by the specific measure of hedonic tone. In Experiment 1, we studied three picture categories with similar affective and semantic contents: 96 affective environmental scenes, which were also converted into 96 cartoons, and 96 representational paintings. Complexity varied along the dimension of elements. In a between-subjects design, each stimulus was presented for 5 s to 206 female participants. Subjective ratings of hedonic tone (either beauty, pleasantness or liking), arousal, complexity and familiarity were collected in three conditions per stimulus set. Complexity and arousal were positively associated in all conditions, with the strongest association observed for paintings. For environmental scenes and cartoons, there was no significant association between complexity and hedonic tone, and the three measures of hedonic tone were highly correlated (all $r_S > 0.85$). As predicted, in paintings the measures of hedonic tone were less strongly correlated (all $r_S > 0.73$), and when controlling for familiarity, the association with complexity was significantly positive for beauty ($r_S = 0.26$), weakly negative for pleasantness ($r_S = -0.16$) and not present for liking. Experiment 2 followed a similar approach and 77 female participants, all non-musicians, rated 92 musical excerpts (15 s) in three conditions of hedonic tone (either beauty, pleasantness or liking). Results indicated a strong relationship between complexity and arousal (all $r_S > 0.85$). When controlling for familiarity effects, the relationship between complexity and beauty followed an inverted-U curve, whereas the relationship between complexity and pleasantness was negative ($r_S = -0.26$) and the one between complexity and liking positive ($r_S = 0.29$). We relate our results to Berlyne's theory and the latest findings in neuroaesthetics, proposing that future studies need to acknowledge the multifaceted nature of hedonic tone in esthetic experiences of artforms.

Keywords: esthetic experience, neuroaesthetics, beauty, pleasantness, liking, complexity, visual arts

OPEN ACCESS

Edited by:
Arthur M. Jacobs,
Free University of Berlin, Germany

Reviewed by:
Enric Munar,
University of the Balearic Islands, Spain
Tomohiro Ishizu,
University College London, UK

****Correspondence:***
Manuela M. Marin
manuela.marin@univie.ac.at;
manuela.marin@uibk.ac.at

Received: *14 July 2016*
Accepted: *10 October 2016*
Published: *04 November 2016*

Citation:
Marin MM, Lampatz A, Wandl M and Leder H (2016) Berlyne Revisited: Evidence for the Multifaceted Nature of Hedonic Tone in the Appreciation of Paintings and Music. Front. Hum. Neurosci. 10:536. doi: 10.3389/fnhum.2016.00536

INTRODUCTION

Esthetic experiences are phenomena that are inherently challenging to study due to a myriad of underlying factors. Therefore, it is not surprising that the field of empirical esthetics has generated competing accounts for why humans show preferences for certain objects over others or why they consider some objects as more beautiful (Palmer et al., 2013; Pelowski et al., 2016). For example, the mere exposure effect (Zajonc, 1968) suggests that the higher the frequency of exposure, the more the stimulus is liked, whereas Berlyne's psychobiological model (Berlyne, 1971) emphasizes the mediating role of arousal in the relationship between stimulus properties and measures of hedonic value (e.g., esthetic pleasure or preference) or interest. Alternatively, Martindale and Moore (1988) and Martindale et al. (1990) propose that more prototypical exemplars of a category are preferred to non-prototypical ones. Fluency theory (Reber et al., 2004; Reber, 2012), which posits that the ease of processing determines preference, may also explain the underlying mechanism of the mere exposure effect and the prototype-preference theory, but it does not offer a clear account for Berlyne's arousal theory (Palmer et al., 2013). Lately, the field has witnessed the emergence of several multicomponent models (Leder et al., 2004; Hagtvedt et al., 2008; Marković, 2012; Shimamura, 2012; Tinio, 2013; Chatterjee and Vartanian, 2014; Leder and Nadal, 2014; Redies, 2015; Pelowski et al., 2016), which aim to describe the essence of esthetic experiences by modeling the relationship between bottom–up and top–down processes as well as their underlying causes and neural substrates. Here, we follow this approach with a strong emphasis on Berlyne's theory.

The field of empirical esthetics has generated a growing body of literature, including divergent results. To be sure, Berlyne's (1971, 1974a) psychobiological model of esthetic experience and his *New Experimental Esthetics* have been very influential in the field (Konecni, 1978; Silvia, 2005), but besides empirical support for his theory a non-negligible number of conflicting results have also been reported (Vitz, 1964; Osborne and Farley, 1970; Kreitler et al., 1974; Nicki and Gale, 1977; Krupinski and Locher, 1988; Neperud and Marschalek, 1988; Messinger, 1998; Stamps Iii, 2002; Eisenberg and Thompson, 2003; Nadal et al., 2010; Marin and Leder, 2013). Therefore, the current research project sought to contribute to the emerging literature by offering explanations for the discrepant research findings regarding Berlyne's theory. For instance, contributions of the multidimensionality of visual complexity (Nadal et al., 2010), individual differences in esthetic preferences (Güçlütürk et al., 2016) as well as a dual-process perspective on preference formation (Graf and Landwehr, 2015) have recently been acknowledged. In this study, we aim to explore alternative pathways in the elucidation of Berlyne's psychobiological model: First, by following a stringent comparative approach, we study the nature of hedonic tone and its relationship with complexity in the experience of three different sets of affective pictures varying in their esthetic quality. Second, we also study the nature of hedonic tone in the appreciation of music using a similar approach. Third, we link our findings to current research in the field of neuroaesthetics.

Being interested in curiosity and exploratory behavior, Berlyne follows an information-theoretic approach and regards the structural features of an artwork as a primary source of esthetic response. To be specific, his psychobiological model is based on Wundt's (1874) curve and suggests that a set of psychophysical, ecological, and collative stimulus features have a certain arousal potential, which in turn determines hedonic value. Very low and high arousal potential leads to low levels of hedonic value, whereas moderate levels lead to high hedonic value, yielding an inverted-U curve which is generated by interacting arousal systems in the brain. According to Berlyne, collative variables, such as novelty, complexity, uncertainty and conflict, are those with the highest arousal potential in the perceiver. The term "hedonic value" comprises several distinct variables, such as pleasure (pleasantness), preference, utility, which can be measured by verbal ratings, as well as reward value and incentive value, which can be measured by non-verbal behavior. A related attribute is beauty (Berlyne, 1971), whose relationship with collative variables has also been frequently studied (Nadal et al., 2010). Berlyne (1974a) later introduced the term "hedonic tone" to refer to verbal expressions of pleasure and the like. Importantly, Berlyne (1974a) states that these measures of hedonic value may or may not have the same underlying psychophysiological basis. Therefore, our findings will have important implications for a better understanding of the neural underpinnings of esthetic experiences (Pearce et al., 2016).

Berlyne's theory has been put to test with a wide range of stimuli, ranging from abstract patterns to paintings and from simple melodic sequences to real music (for a review see (Nadal, 2007; Nadal et al., 2010; Marin and Leder, 2013). Although Berlyne's psychobiological model has been fruitful in the study of psychology and the arts (Silvia, 2005; Jacobsen, 2006), it has also generated criticism, mostly stemming from the fact that the inverted-U shape relationship between collative variables and hedonic value could not always be replicated. For example, Cupchik (1988) regretted that Berlyne's skepticism about cognitive psychology hindered the study of how thought and emotion interact during esthetic experiences, which implies a disinterest in the study of understanding, a critical attribute of art perception (Leder et al., 2004). Konecni (1996) criticized Berlyne for not paying enough attention to sympathetic arousal and emotion as well as to the context of esthetic behavior. Similarly, Silvia (2005) pointed out that an arousal theory cannot account for the diversity of emotions experienced during art perception. Furthermore, the concept of arousal that Berlyne had in mind has not been supported by neurophysiological research (Nicki, 1972; Nicki and Gale, 1977; Silvia, 2005; Nadal, 2007), and others proposed that prototypicality is better able to explain esthetic preference than a collative motivation model (Whitfield, 1983; Martindale and Moore, 1988). Last but not least, it has been suggested that single-factor explanations, in this case the modulation of arousal through a set of stimulus features, may not be sufficient to capture the nature of esthetic experiences (Leder and Nadal, 2014).

However, these objections do not justify abandoning Berlyne's theory, especially since there is cumulative evidence that arousal plays an important role, not only in emotions in general (Russell,

1980), but specifically in esthetic experiences. First, it should be pointed out that discrepancies regarding Berlyne's theory seem mostly relate to the inverted-U shape relationship between collative variables and hedonic value, and not to the predicted linear relationship between collative variables and subjective arousal. Even in studies that could not confirm the inverted-U shape relationship, a positive relationship between collative variables and subjective arousal was reported when artistic (Krupinski and Locher, 1988; Neperud and Marschalek, 1988; Marin and Leder, 2013) and non-artistic stimuli (Heath et al., 2000; Marin and Leder, 2013) were investigated. This suggests that a differentiated view on Berlyne's theory is required. Second, arousal has recently been identified as a mediator between the relationship of complexity and pleasantness for visual stimuli, and to a lesser degree, for music (Marin and Leder, 2013, 2016). Third, there is a large body of empirical evidence, not only in light of Berlyne's model, suggesting a relationship between arousal and stimulus features in the visual as well as auditory domain (Jacobs and Hustmyer, 1974; Bertamini et al., 2013; Marin and Leder, 2013; Gingras et al., 2014). Fourth, a link between arousal and preference has frequently been reported (Konecni et al., 1976; Konecni and Sargent-Pollock, 1977; North and Hargreaves, 2000; Schäfer and Sedlmeier, 2011; Blijlevens et al., 2012; Ramsøy et al., 2012). Bearing in mind that arousal may generally affect judgment, learning and memory (Storbeck and Clore, 2008), we thus propose that the concept of arousal should be more systematically studied in the context of Berlyne's theory, especially since alternative theories, such as appraisal theory (Silvia, 2005), also include an arousal component.

Research in the line of Berlyne (1974a) has originally followed a program using verbal ratings as well as psychophysiological and behavioral measures as dependent variables. Verbal ratings usually comprised a large set of *descriptive scales* (measuring collative variables), *evaluative scales* (measuring hedonic value and related attributes), *internal-state scales* (measuring arousal, tension and pleasure) as well as, to a lesser degree, *stylistic scales* (referring to stylistic features of an artwork). Factor analysis conducted on these variables often showed pleasure to be correlated with other evaluative scales (see Berlyne, 1974a). Early research in the tradition of Berlyne's *New Experimental Esthetics* thus usually worked with more than one rating scale for each scale category, which has become less usual in recent times. Nowadays, a small set of scales and items is generally used to study esthetic experiences in the context of Berlyne's theory, probably due to the fact that research has shown that Berlyne's large set of scales can be reduced to 2–3 factors (e.g., hedonic tone, arousal, and complexity) (Berlyne, 1974a). Although a reduction in the number of rating scales is generally welcome, especially because it allows for simpler research designs and shorter experiments, many researchers after Berlyne do not explain why they have chosen one rating scale out of several possible ones. In other words, it is often unclear why authors chose to measure beauty instead of liking or pleasantness as a measure of hedonic tone. Therefore, it can be argued that the multifaceted nature of hedonic tone (by which we refer to the many aspects this term originally comprised) has not received enough attention in the field. We believe that it is necessary to acknowledge the concepts behind different measures of hedonic tone and to study their relationships with complexity in different stimulus sets varying in esthetic quality.

The main goal of the current study was to clarify how verbal reports of different measures of hedonic tone relate with visual and musical complexity, aiming to contribute to the elucidation of Berlyne's theory (Nadal et al., 2010; Marin and Leder, 2013; Güçlütürk et al., 2016). Recent empirical work has demonstrated that the relationship between complexity and beauty can be explicated by considering the multidimensionality of complexity (Nadal et al., 2010), which is the finding that the experience of complexity can be produced by variations of (1) the number and diversity of elements, (2) disorganization and (3) asymmetry. Nadal et al. (2010) showed that if complexity is varied along the dimension of elements, a positive linear relationship with beauty is observed. Variation along the dimension of disorganization revealed a U-shape relationship with beauty, whereas indications for an inverted-U shape relationship were found for variation along the dimension of symmetry. We were inspired by this approach and explored the possibility that the specific measure of hedonic tone may determine the relationship with a specific dimension of complexity.

Concepts such as complexity, arousal, pleasantness, liking, and beauty are also essential features of more recent models of esthetic experiences (Leder et al., 2004; Brattico et al., 2013; Leder and Nadal, 2014; Pelowski et al., 2016), thus the study of their interrelationships is of general interest to the field. If it can be shown that complexity and measures of hedonic tone are associated in different ways, and importantly, that measures of hedonic tone are related, the data would directly justify the integration of these concepts into one model of esthetic experience. Importantly, the outcome of the current study may have direct consequences for linking cognitive models of esthetic experience with current neuroimaging data (Cela-Conde et al., 2013; Chatterjee and Vartanian, 2014; Leder et al., 2015). In the last decade, neuroimaging studies on the appreciation of visual art have suggested that, rather than a single brain area, there is a complex neural system that underlies esthetic experiences (Nadal, 2013; Cela-Conde and Ayala, 2014; Vartanian and Skov, 2014; Boccia et al., 2016).

Nadal (2013) suggested that at least three different sets of brain regions underlie esthetic experiences, and these brain regions may be differently involved in the processing of hedonic tone. Evaluative judgments, attentional processing and memory retrieval are related to activations in the prefrontal, parietal and temporal cortical regions. Esthetic experiences also involve the reward circuit as well as low-, mid-, and high-level cortical sensory regions (see also Cela-Conde et al., 2013). In the context of the current study, brain regions related to evaluative judgments and the reward circuit are of particular interest. Activation of the dorsolateral prefrontal cortex (Cela-Conde et al., 2004) and the anterior medial prefrontal cortex (Jacobsen and Höfel, 2003; Jacobsen et al., 2006; Kirk, 2008) may indicate the engagement of evaluative judgment processes, in this case mostly related to beauty. Nadal (2013) suggested that, since these activations occur between 400 and 600 ms after picture onset, they may be involved in the formation of initial impressions which impact

on further processing stages such as attention, perception and response selection (Cela-Conde et al., 2011). Other brain areas reported to be active during evaluative judgments (of beauty or appeal) are the ventrolateral prefrontal cortex (Jacobsen et al., 2006; Kirk, 2008), the left temporal pole (Jacobsen et al., 2006), as well as the posterior cingulate cortex and precuneus (Jacobsen et al., 2006; Kirk et al., 2009).

Several neuroimaging reports on the esthetic experience of art have also confirmed activation of different brain areas constituting the reward circuit (Kawabata and Zeki, 2004; Kirk, 2008; Ishizu and Zeki, 2011; Lacey et al., 2011), which may be related to the experience of pleasure or pleasantness as well as liking responses. The activity of the medial orbitofrontal cortex was positively associated with liking or preference (Kawabata and Zeki, 2004; Kirk, 2008), whereas the activity in the lateral orbitofrontal cortex was negatively associated with appeal ratings (Kirk, 2008; Munar et al., 2012). Another brain area of the reward circuit, the anterior cingulate cortex, was found to be active when people were engaged with an artwork they liked (Vartanian and Goel, 2004; Cupchik et al., 2009; Kirk et al., 2009; Boccia et al., 2015; Yeh et al., 2015). Both the orbitofrontal cortex and the anterior cingulate cortex were active when cognitive and affective processes interacted with each other during the evaluation of sensory information (Pessoa, 2008; Rolls and Grabenhorst, 2008). The insular cortex was more active when viewers experienced emotions induced by an artwork. Moreover, sub-cortical components of the reward circuit, such as the nucleus accumbens, have frequently been reported to be active during pleasurable esthetic experiences (Vartanian and Goel, 2004; Cupchik et al., 2009; Kirk et al., 2009; Ishizu and Zeki, 2011). Taken together, these neuroimaging studies suggest that judgments of beauty, pleasantness and liking may not be necessarily processed by the same set of brain regions. Similar conclusions have also been drawn by Brattico et al. (2013) who modeled the neural correlates of music-induced esthetic experiences.

There is a recent interest in comparing esthetic experiences across object classes and sensory domains (for a review see Marin, 2015), and Berlyne's psychobiological model, regardless of the debate about its underlying mechanisms, may be adopted as a scientific playground to study the nature of esthetic experiences (Normore, 1974; Marin and Leder, 2013). Berlyne's theory made clear predictions and encompasses general concepts that can meaningfully be studied in different sensory domains and in different object classes within one sensory domain. Recently, by studying affective environmental scenes, representational paintings and music, we found that the relationship between complexity and arousal seems to be stronger in the musical than in the visual domain (Marin and Leder, 2013). We showed that mediation effects of arousal may be more prominent in the visual than in the musical domain, and further, that objective measures of visual complexity can be used in their analogous forms for studying musical complexity. Regarding the relationship between hedonic tone (pleasantness) and complexity, a complex interplay between familiarity, participant gender and stimulus type determined the kind of relationship. Neither the predictions made by Berlyne (1971) (inverted-U shape) nor those of Nadal et al. (2010) (linear positive for variations of the number of elements) were confirmed. Since pleasantness, the measure of hedonic tone, was also decisive in the pre-selection of the affective stimuli, we argued that other measures of hedonic tone should be used to study the relationship with complexity (Marin and Leder, 2013). In other words, not only the multidimensional nature of complexity may play a role in the relationship with hedonic tone but also the type of hedonic measure (e.g., pleasantness, liking, beauty etc.).

The results of previous studies using more than one measure of hedonic tone in the visual domain indeed suggest that the interrelationships between them may vary according to the stimulus type. For example, Berlyne and Ogilvie (1974) asked participants to rate a set of paintings stemming from the 14th to the 20th century on 12 ratings scales, including beauty, pleasantness, and pleasure. Importantly, they found that these scales, although loading on the same hedonic tone factor, correlated somewhat differently with the complexity scale in their two small groups of participants, making it difficult to draw any conclusions. Normore (1974) studied animations of a dot whose dynamics regarding position, brightness and duration varied in complexity. In comparison to Berlyne and Ogilvie (1974) study, ratings of beauty and pleasingness were nearly identical. Berlyne (1974b) also investigated visual patterns and, among other things, made participants rate beauty, pleasingness and liking. He observed that beauty and pleasingness were highly positively correlated, which also held true for the relationship between beauty and liking as well as for the relationship between pleasingness and liking. In a between-subjects design, Russell and George (1990) further investigated evaluative responses to 15 paintings of diverse Western styles by obtaining ratings of preferability, pleasingness, likability, complexity, and familiarity. The negative correlation between complexity and pleasingness was stronger than those correlations observed for preferability and likability. Likability and preferability as well as likability and pleasingness were highly correlated, whereas the relationship between pleasingness and preferability was slightly weaker, but these correlations were not corrected for familiarity effects. Looking at architectural stimuli, Imamoglu (2000) investigated esthetic responses to different types of house façade drawings and collected ratings of pleasantness, beauty and liking. The results indicated that the three measures of hedonic tone were highly correlated. Sato and Oda (2014) studied the effects of shape and color on the esthetic evaluation of colored shapes varying in complexity and curvature. The results showed that shape and color affected beauty, pleasantness and liking ratings slightly differently. In summary, these findings suggest that the interrelationship between the three measures of hedonic tone may be stronger for naturalistic pictures and abstract patterns than for artworks.

In the current study, we followed a systematic comparative approach – across stimulus categories – and sought to examine (i) the nature of hedonic tone and its relationship with visual complexity as defined by the number and diversity of elements, and (ii) the nature of hedonic tone and its relationship with musical complexity. We studied three evaluative scales frequently used as measures of hedonic tone, namely pleasantness,

beauty and liking. In Experiment 1, we collected subjective reports on affective environmental scenes (IAPS pictures) and representational paintings (Marin and Leder, 2013), in which complexity was mostly varied along the dimension of elements (Marin and Leder, 2016). Environmental scenes were photographs of everyday-life scenes, involving a range of semantic categories and various emotional contents, which are widely used to study emotion and attention. Representational paintings, mostly stemming from the 19th century, also comprised a wide range of semantic and emotional contents but were not matched on a one-to-one basis with environmental scenes. Although these two picture sets were similar in semantic and emotional contents, and thus allowed for some conclusions about the nature of esthetic experiences, we also introduced a new, third picture set, namely environmental scenes transformed into cartoons by means of photo editing software. These cartoon-like pictures thus contain identical semantic content, preserve composition and ordering of objects, but also render a visually poorer, reduced representation, and lack the distinctive visual features that mark an artist's style. This enabled us to compare two picture sets that differed in their artistic quality but not necessarily in their emotional and semantic contents. Moreover, the cartoons represented a uniform "artistic" style, which was not exactly the case for representational paintings. In Experiment 2, we used musical stimuli taken from Experiment 3 as described in Marin and Leder (2013) and followed a similar procedure as in Experiment 1.

We decided to employ a between-subjects design in which participants rated pictures or musical excerpts for familiarity, complexity, and arousal as well as for only one of the three measures of hedonic tone. This ensured that participants could not guess the aim of the experiment, and further, this design was also closer to what other researchers mostly employed in their designs, i.e., to use only one measure of hedonic tone. Last, we decided to restrict our sample to females because gender effects regarding these picture sets have been reported earlier (Marin and Leder, 2013). We also controlled for mood prior to the experiment (Flexas et al., 2013; Gartus and Leder, 2014) as well as for interest in visual arts (Experiment 1) and musical sophistication (Experiment 2).

For Experiment 1 we hypothesized that, in line with Nadal et al. (2010), beauty would show a positive association with complexity as defined by the number and diversity of elements. Based on Marin and Leder (2013, 2016), we predicted that there would be no significant relationship between this complexity dimension and pleasantness in environmental scenes and paintings if familiarity is controlled for. Regarding liking ratings and the relationship with complexity it was difficult to make a concrete prediction, except for arguing for an inverted-U shape relationship, following Berlyne's theory. In addition, we surmised that the relationship between complexity and arousal would be positive in all picture sets but stronger in paintings than in environmental scenes and cartoons (Marin and Leder, 2013). Furthermore, we hypothesized that the multifaceted nature of hedonic tone in relation to complexity may be most obvious regarding visual arts in comparison to environmental scenes and cartoons (Berlyne and Ogilvie, 1974; Russell and George, 1990; Imamoglu, 2000).

For Experiment 2, our hypotheses were less concrete because much less is known about the underlying dimensions of musical complexity. However, as objective measures of musical complexity have indicated in Marin and Leder (2013), event density moderately correlated with ratings of subjective complexity. Based on this finding and findings by Nadal et al. (2010), we predicted a positive association between beauty and complexity, assuming that complexity and beauty are similarly related in both the visual and musical domains. Following findings by Marin and Leder (2013) for the group of females, we predicted no significant association between pleasantness and complexity after controlling for familiarity effects. For the complexity-liking association we were only able to refer to Berlyne's original theory and his predicted inverted U-shape relationship. Furthermore, we predicted a strong relationship between complexity and arousal in all experimental conditions.

EXPERIMENT 1

Materials and Methods

Participants

Two-hundred thirty female psychology students at the University of Vienna participated in the experiment in exchange for course credit. All participants were unfamiliar with the three picture sets and did not participate in previous studies by Marin and Leder (2013, 2016). Participants had normal or corrected-to-normal visual acuity and normal color vision. The participants were randomly assigned to one of nine groups that differed regarding the three picture sets and the type of hedonic measure to be rated (beauty, pleasantness and liking).

The participants of the nine groups were screened for outliers with respect to their mood prior to the experiment (participants with very low mood scores were removed), their art interest as well as the trait emotional intelligence, empathy and stress reactivity scales (SRS; several participants scored very low on the empathy and emotional intelligence scales). Twenty-four participants were removed for these reasons by investigating the respective boxplots (an outlier was defined as a data point more than 1.5 interquartile ranges below the first quartile or above the third quartile). The remaining 206 participants in the nine groups did not significantly differ regarding their age, mood prior to the experiment and their scores on the trait EI, empathy quotient and SRS. We also ensured that art interest was similar, and generally low, in the six groups rating cartoons and paintings (see **Table 1**).

Materials

The 96 affective environmental scenes [taken from the International Affective Picture System (IAPS), Lang et al., 2005], and 96 representational paintings used in this experiment were identical to those used in Marin and Leder (2013, 2016). The pictures were in landscape format and varied in their semantic and emotional contents as well as in their complexity (figure-ground composition vs. complex scene). There were no pictures containing mutilation, erotic scenes or brand names. Marin and

TABLE 1 | Participant characteristics.

Group	Age [years]	Pos./neg. mood A	Alertness/ fatigue A	Quietude/ disquietude A	Art interest	Trait EI	EQ short	SRS
IAPS Beauty	*M* 23.8	17.29	13.00	15.08	–	156.88	70.63	58.13
(*n* = 24)	*SD* 4.0	1.81	2.95	3.31		16.39	7.30	8.72
IAPS Pleasantness	*M* 22.1	17.57	13.14	16.43	–	163.52	71.67	57.00
(*n* = 21)	*SD* 2.9	2.18	3.81	2.62		10.41	5.03	8.46
IAPS Liking	*M* 22.2	16.22	12.65	15.26	–	151.17	70.61	59.70
(*n* = 23)	*SD* 3.3	3.01	2.99	2.49		13.23	5.65	9.76
Cartoons Beauty	*M* 24.3	17.04	14.44	14.96	63.40	153.48	71.04	56.96
(*n* = 25)	*SD* 5.3	2.54	3.83	3.18	16.64	18.12	6.41	8.17
Cartoons Pleasantness	*M* 21.8	16.64	12.73	14.82	64.05	157.50	72.95	59.72
(*n* = 22)	*SD* 1.9	2.50	3.40	2.46	11.94	12.69	5.07	9.32
Cartoons Liking	*M* 22.6	17.29	14.04	16.13	59.38	160.46	72.42	57.38
(*n* = 24)	*SD* 2.4	1.94	3.30	2.76	15.10	13.70	5.91	8.84
Paintings Beauty	*M* 22.6	16.41	12.77	15.18	62.00	158.45	73.14	61.73
(*n* = 22)	*SD* 4.0	2.58	3.57	3.06	14.40	15.85	6.29	7.75
Paintings Pleasantness	*M* 22.4	17.35	13.30	15.40	66.95	154.65	70.90	60.95
(*n* = 20)	*SD* 4.0	1.84	2.66	2.54	11.17	14.38	5.92	8.13
Paintings Liking	*M* 22.2	16.92	13.64	14.84	61.24	153.04	69.64	57.96
(*n* = 25)	*SD* 3.1	1.89	4.28	2.51	15.03	22.77	7.30	8.87
Test statistic	*H* = 10.49	*H* = 5.55	*H* = 6.59	*H* = 7.87	*F* = 0.73	*F* = 1.38	*F* = 0.85	*F* = 0.91
p	0.232	0.698	0.581	0.447	0.605	0.207	0.560	0.513
df	8	8	8	8	5,137	8,197	8,197	8,197

Summary of the nine groups' mean age, mood prior to the experiment (MDBF short form A), art interest, trait emotional intelligence, empathy quotient and stress reactivity. M, mean; SD, standard deviation; n, number of participants; p, calculated probability; df, degrees of freedom; H, Kruskal–Wallis test statistic; F, F-test statistic; IAPS, International Affective Picture System; pos., positive; neg., negative; EI, emotional intelligence; EQ, empathy quotient; SRS, stress reactivity scale. Kruskal–Wallis tests conducted for the three conditions per picture set revealed no significant differences, except for a significant difference of trait EI between the three IAPS conditions, H(2) = 9.56, p = 0.008.

FIGURE 1 | Illustration of stimulus categories employed in Experiment 1. (A) Photograph of a landscape (Manuela Marin), **(B)** photograph converted into a cartoon-like picture, and **(C)** representational painting by Paul Cézanne, *The bay of Marseille, seen from L'Estaque* (1885). Note that pictures **(A,B)** were not used in the actual experiment but are depicted for illustrative purposes only.

Leder (2016) showed that the subjective complexity ratings of these pictures are mostly determined by the number of objects, and to a lesser degree, by their disorganization. All scenes were converted into cartoon-like pictures using Adobe Photoshop CS5 and its filter functions (**Figure 1**). The first step involved the following processing stages: Filter-Artistic-Poster Edges (edge thickness = 0, edge intensity = 0, posterization = 1). Next, the Cutout filter was applied from the Photoshop filter gallery and the settings for number of levels, edge simplicity and edge fidelity were individually adjusted to obtain a picture whose semantic content could be easily deduced. In other words, the level of abstraction was not very high. Last, the brightness and contrast settings were adjusted (Image, Adjustment, brightness = 10, contrast = 30).

Standardized self-report measures included the two short forms A and B of the multidimensional mood questionnaire (MDBF, Steyer et al., 1997), the short trait emotional intelligence questionnaire (TEIQue-SF, Petrides and Furnham, 2006), the stress reactivity scale (SRS, Schulz et al., 2005) and a German version of the short Empathy Quotient (EQ-Short, Wakabayashi et al., 2006). For the groups rating representational paintings and cartoons, a short 15-item unpublished questionnaire on interest in visual arts was used. It comprised statements (developed in our research group), such as, "I enjoyed arts education at school," "I am interested in art," and "I don't like ugly artworks," and participants indicated their answers on 7-point scales ranging from "fully disagree" to "fully agree" (see Supplementary Table).

A self-developed questionnaire was administered after the performance of the experimental task. Participants were asked to report, on 7-point scales, how much they liked the pictures in general; how many they have seen before; whether they have experienced feelings while watching the pictures, and how strong these feelings were; whether they were emotionally engaged with the contents of the pictures; and how difficult it was to judge complexity. In the three conditions involving cartoons, participants also reported whether they would consider these pictures as art.

Procedure

We followed the current ethical guidelines at the University of Vienna and the version of the Ethical Principles of Psychologists and Code of Conduct of the American Psychological Association (with 2010 amendments). All participants read and signed an informed consent form prior to the experiment.

In general, the experimental procedure followed the one described in Experiments 1 and 2 in Marin and Leder (2013) and the one described in Experiment 1 in Marin and Leder (2016), but the order of the rating scales differed and standardized questionnaires were added. There were nine groups, which were defined by the type of picture set (environmental scenes, cartoons, and representational paintings) and the type of hedonic tone (beauty, pleasantness and liking). Participants filled the short form A of the multidimensional mood questionnaire before the actual experiment began. Then, sitting 60–70 cm away from the screen (19-inch Iiyama ProLite B1906S), participants were familiarized with the experimental task in two practice trials. Each trial was announced for 5 s and then a picture was shown on a black background for 5 s, after which the first (out of four) 7-point rating scale appeared on the screen, i.e., the one referring to hedonic tone. Ratings were given by mouse click and self-paced. Next, ratings of subjective arousal, complexity and familiarity were collected, after which the following trial was announced on the screen. Participants were asked to look at the picture for the entire presentation duration, and the picture was not shown when the ratings were collected. All pictures were blocked according to emotional contents (low-arousing pleasant, low-arousing unpleasant, high-arousing pleasant and high-arousing unpleasant) as pre-selected by Marin and Leder (2013), with an equal number of 23 trials in each block and a total number of 96 trials in each condition. The orders of the four blocks as well as the order of the pictures within each block were randomized. The participants were told that the blocks would vary in their emotional content, and they were allowed to take breaks between blocks.

The instructions for the measures of hedonic tone were as follows (note that only one of the scales was presented in each group of participants): Pleasantness: "Please rate the degree of pleasantness of your emotional experience," with (1) "very unpleasant" and (7) "very pleasant" as anchors; Beauty: "Please report the degree of perceived beauty of the picture" with (1) "not at all beautiful" and (7) "very beautiful" as anchors; Liking: "Please report how much you liked the picture" with (1) "not at all" and (7) "very much" as anchors. Instructions for arousal ratings were "Please rate your felt arousal," with (1) "very calm" and (7) "very excited" as anchors. Complexity was assessed by the following instruction: "Please rate your felt complexity of the picture," with (1) "very simple" and (7) "very complex" as anchors. Finally, participants rated their familiarity with the picture content: "Please rate your familiarity with the contents of the picture," with (1) "unfamiliar" and (7) "very familiar" as anchors.

After participants had finished the experiment, they filled the short form B of the multidimensional mood questionnaire, the self-developed questionnaire on the experiment, followed by the TeiQue-SF, SRS and EQ-short scales, and finally, the art interest scale (in the groups ratings cartoons and artworks). The experimental sessions lasted around 60 min. Participants were thanked, debriefed and dismissed.

Statistical Analysis

The analytical procedure regarding the analysis was identical to Marin and Leder (2013, 2016), thus enabling a meaningful comparison of the results across studies. The picture was chosen as the unit of analysis, which implies that the results need to be interpreted with regard to pictures and not participants. The main analysis was run using IBM SPSS Statistics 21. All statistical tests were computed at an alpha level of 0.05 and two-tailed. The Bonferroni–Holm procedure (Holm, 1979) was used to control the family wise error rate and computed in Matlab 2014b (The MathWorks Inc., Natick, MA, USA).

Results

In a first step, the subjective ratings were averaged across participants for each picture in each condition. In a second step, the inter-rater reliability was assessed by calculating the average intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency (Shrout and Fleiss, 1979; McGraw and Wong, 1996). **Table 2** shows that the ICCs were generally high (all ICCs > 0.7), which justifies averaging across participants and considering pictures as unit of analysis. In a third step, outliers were determined by examining boxplots of each type of rating for each category of pictures. There were no outliers in the sets of IAPS pictures and cartoons. Familiarity and beauty ratings of paintings revealed two outliers, which were removed. Since most distributions were not normal, non-parametric analyses were employed to investigate

TABLE 2 | Intra-class correlation coefficients (ICCs) for subjective ratings of familiarity, complexity, arousal and hedonic tone (beauty, pleasantness and liking).

Group	ICC(2,k) Familiarity	ICC (2,k) Complexity	ICC (2,k) Arousal	ICC (2,k) Hedonic tone
IAPS Beauty ($n = 24$)	0.934 CI [0.914,0.952]	0.875 [0.836,0.909]	0.869 [0.828,0.904]	0.961 [0.948,0.971]
IAPS Pleasantness ($n = 21$)	0.926 CI [0.903,0.946]	0.845 [0.797,0.887]	0.859 [0.815,0.897]	0.976 [0.968,0.982]
IAPS Liking ($n = 23$)	0.937 CI [0.917,0.954]	0.884 [0.848,0.915]	0.900 [0.868,0.927]	0.963 [0.952,0.973]
Cartoons Beauty ($n = 25$)	0.948 CI [0.932,0.962]	0.819 [0.762,0.867]	0.862 [0.818,0.899]	0.927 [0.904,0.946]
Cartoons Pleasantness ($n = 22$)	0.927 CI [0.904,0.947]	0.836 [0.785,0.880]	0.852 [0.805,0.891]	0.969 [0.960,0.978]
Cartoons Liking ($n = 24$)	0.941 CI [0.922,0.956]	0.818 [0.761,0.867]	0.887 [0.851,0.917]	0.957 [0.943,0.968]
Paintings Beauty ($n = 22$)	0.764 CI [0.690,0.828]	0.836 [0.785,0.880]	0.824 [0.768,0.871]	0.891 [0.856,0.920]
Paintings Pleasantness ($n = 20$)	0.765 CI [0.691,0.828]	0.867 [0.826,0.903]	0.814 [0.756,0.864]	0.964 [0.952,0.973]
Paintings Liking ($n = 25$)	0.788 CI [0.722,0.845]	0.913 [0.886,0.937]	0.858 [0.814,0.896]	0.918 [0.892,0.940]

ICC(2,k), two-way random average measure, type consistency; CI, 95% confidence interval. ICCs were calculated for ratings of 96 pictures in each category (IAPS, cartoons and paintings).

the interrelationships between the four types of ratings in each condition.

Table 3 depicts the interrelationships between the four types of ratings, separately given for each of the nine conditions. Note that all relationships were visually inspected to ensure that they were monotonous. Familiarity with IAPS pictures and cartoons correlated negatively with complexity and arousal [ranging from small (r_s ~0.2) to medium (r_s ~0.3) effect sizes], with stronger relationships for cartoons than IAPS pictures. The relationships between familiarity and measures of hedonic tone were positive, strong, and constant across the three measures of hedonic tone ($r_s > 0.6$), and comparable in IAPS pictures and cartoons. A similar pattern of results was observed for paintings; however, there was more variation across the three measures of hedonic tone in terms of effect size.

Controlling for familiarity effects, **Table 4** shows how complexity, arousal and the three measures of hedonic tone were related. In line with Berlyne's prediction, we generally observed small to moderate associations between complexity and arousal for both IAPS pictures and cartoons. In paintings, this association was somewhat stronger (r_s ~0.5), nearly reaching the benchmark of a large effect (Cohen, 1988), supporting our prediction. Furthermore, the data suggested that the relationship between complexity and arousal was the strongest in the liking condition, regardless of the picture type.

Of particular interest was the comparison of the relationships between complexity and the respective measures of hedonic tone across the three picture sets. **Figures 2–4 (A–C)** show these relationships including a regression line based on results of curve-fitting analyses conducted in SigmaPlot 13.0. By not controlling for familiarity effects, the multifaceted nature of hedonic tone became visually apparent in cartoons and paintings but not in IAPS pictures. The relationship between complexity and pleasantness was negative in cartoons, whereas there were no relationships for the beauty and liking conditions. In paintings, the relationships differed to a much larger extent and the graphs seemed to indicate a positive association between complexity and beauty, a negative association between complexity and pleasantness, and no relationship in the liking condition.

Turning to the results of partial correlations controlling for familiarity effects, we predicted a positive relationship for beauty based on results by Nadal et al. (2010). In IAPS pictures and cartoons we found weak indications for a positive association, but these relationships were not significant (**Table 4**). The relationship between complexity and beauty was positive and significant only for paintings ($r_s = 0.26$). Next, we predicted no relationship between complexity and pleasantness, which held true for IAPS pictures. For cartoons and paintings we observed indications of a weak, non-significant negative relationship. The relationships between complexity and liking ratings were absent in IAPS pictures and paintings, but weakly non-significantly positive in cartoons. We also compared the correlation coefficients expressing the relationship between complexity and hedonic tone by means of a Fisher r-to-z transformation in paintings. The coefficients of the beauty-complexity and pleasantness-complexity relationships significantly differed from

TABLE 3 | Spearman's rank order correlations between measures of familiarity, complexity, arousal and measures of hedonic tone in response to IAPS pictures, cartoons and paintings.

Stimulus type	Group	Measure	Familiarity	Complexity
IAPS	Beauty	Complexity	−0.19	
	Pleasantness		−0.19	
	Liking		−0.29*	
	Beauty	Hedonic tone	0.66*	−0.05
	Pleasantness		0.62*	−0.15
	Liking		0.62*	−0.24*
N = 96	Beauty	Arousal	−0.30*	0.34*
df = 94	Pleasantness		−0.24	0.29*
	Liking		−0.19	0.40*
Cartoons	Beauty	Complexity	−0.33*	
	Pleasantness		−0.31*	
	Liking		−0.36*	
	Beauty	Hedonic tone	0.68*	−0.13
	Pleasantness		0.61*	−0.29*
	Liking		0.62*	−0.10
N = 96	Beauty	Arousal	−0.44*	0.34*
df = 94	Pleasantness		−0.27*	0.30*
	Liking		−0.34*	0.45*
Paintings	Beauty	Complexity	0.09	
	Pleasantness		−0.16	
	Liking		0.11	
	Beauty	Hedonic tone	0.64*	0.25*
	Pleasantness		0.71*	−0.22
	Liking		0.59*	0.08
N = 94	Beauty	Arousal	−0.28*	0.42*
df = 92	Pleasantness		−0.52*	0.47*
	Liking		−0.29*	0.43*

$p < 0.05$ after Bonferroni–Holm correction; df, degrees of freedom.

TABLE 4 | Partial Spearman's rank order correlations controlling for familiarity between measures of complexity, arousal and hedonic tone in response to IAPS pictures, cartoons and paintings.

Stimulus type	Group	Measure	Complexity	Hedonic tone
IAPS	Beauty	Hedonic tone	0.10	
	Pleasantness		−0.04	
	Liking		−0.08	
N = 96	Beauty	Arousal	0.30**	N/A
df = 93	Pleasantness		0.26*	N/A
	Liking		0.37***	N/A
Cartoons	Beauty	Hedonic tone	0.13	
	Pleasantness		−0.14	
	Liking		0.16	
N = 96	Beauty	Arousal	0.23*	N/A
df = 93	Pleasantness		0.24*	N/A
	Liking		0.37***	N/A
Paintings	Beauty	Hedonic tone	0.26*	
	Pleasantness		−0.16	
	Liking		0.02	
N = 94	Beauty	Arousal	0.46***	−0.07
df = 91	Pleasantness		0.46***	−0.56***
	Liking		0.49***	−0.23*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; df, degrees of freedom. The relationships between the respective measures of hedonic tone and arousal were quadratic in IAPS pictures and cartoons and are thus not reported.*

each other ($p = 0.004$), whereas the beauty-complexity and liking-complexity relationships ($p = 0.097$) as well as the pleasantness-complexity and liking-complexity relationships ($p = 0.223$) were not significantly different form each other. Taken together, these results show that the relationship between complexity and hedonic tone depends on the specification of the latter concept (i.e., beauty, pleasantness and liking), which became mostly apparent in the study of paintings.

Due to the pre-selection of the stimuli, the relationships between the respective measures of hedonic tone and arousal were quadratic in IAPS pictures (see Marin and Leder, 2013, 2016) and cartoons and did not change much across the three conditions of hedonic tone. However, two of these relationships were linear in paintings (**Table 4**): we observed negative relationships for pleasantness ($r_s = -0.56$) and liking ($r_s = -0.23$) but not for beauty, providing more evidence for the multifaceted nature of hedonic tone in the perception of paintings.

If hedonic tone is not a uniform concept when considering the perception of paintings, intercorrelations between the different measures of hedonic tone should be lower for ratings of paintings than for IAPS pictures and cartoons. In IAPS pictures ($N = 96$), the correlation strengths between measures of hedonic tone were very high: $r_{\text{beauty-pleasantness}} = 0.91$, $r_{\text{beauty-liking}} = 0.95$, and $r_{\text{liking-pleasantness}} = 0.92$. In cartoons ($N = 96$), these relationships were similar, but the beauty-pleasantness relationship was slightly lower: $r_{\text{beauty-pleasantness}} = 0.85$, $r_{\text{beauty-liking}} = 0.90$, and $r_{\text{liking-pleasantness}} = 0.92$. However, in paintings ($N = 94$), the three measures correlated to a lesser extent with each other: $r_{\text{beauty-pleasantness}} = 0.73$, $r_{\text{beauty-liking}} = 0.91$, and $r_{\text{liking-pleasantness}} = 0.77$. Overall the pattern of results suggests that the relationship between beauty and pleasantness as well as the one between liking and pleasantness were weaker for paintings than for the other two stimulus sets.

Intercorrelations between the three complexity ratings of each picture set revealed that they were generally highly correlated (subscripts refer to the three conditions per picture set): IAPS pictures: $r_{\text{beauty-pleasantness}} = 0.90$, $r_{\text{beauty-liking}} = 0.85$, and $r_{\text{liking-pleasantness}} = 0.87$; cartoons: $r_{\text{beauty-pleasantness}} = 0.84$, $r_{\text{beauty-liking}} = 0.87$, and $r_{\text{liking-pleasantness}} = 0.80$; paintings: $r_{\text{beauty-pleasantness}} = 0.85$, $r_{\text{beauty-liking}} = 0.85$, and $r_{\text{liking-pleasantness}} = 0.88$. For arousal ratings, these intercorrelations across the three conditions per picture set were of similar strength: IAPS pictures: $r_{\text{beauty-pleasantness}} = 0.85$, $r_{\text{beauty-liking}} = 0.90$, and $r_{\text{liking-pleasantness}} = 0.87$; cartoons: $r_{\text{beauty-pleasantness}} = 0.81$, $r_{\text{beauty-liking}} = 0.87$, and $r_{\text{liking-pleasantness}} = 0.82$; paintings: $r_{\text{beauty-pleasantness}} = 0.81$, $r_{\text{beauty-liking}} = 0.83$, and $r_{\text{liking-pleasantness}} = 0.85$.

The mean complexity ratings of all pictures per type of category and condition were as follows: IAPS pictures: $M_{\text{beauty}} = 3.87$, $SD_{\text{beauty}} = 0.77$, $M_{\text{pleasantness}} = 3.65$,

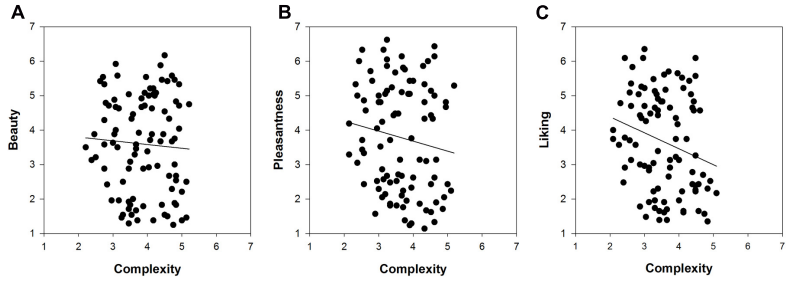

FIGURE 2 | The relationships between complexity and the three measures of hedonic tone for 96 IAPS pictures: (A) beauty, (B) pleasantness, and (C) liking.

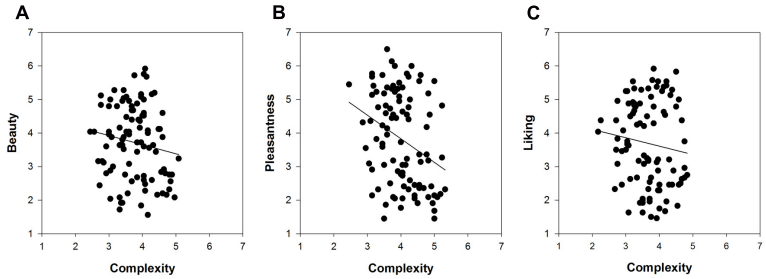

FIGURE 3 | The relationships between complexity and the three measures of hedonic tone for 96 cartoons based on IAPS pictures: (A) beauty, (B) pleasantness, and (C) liking.

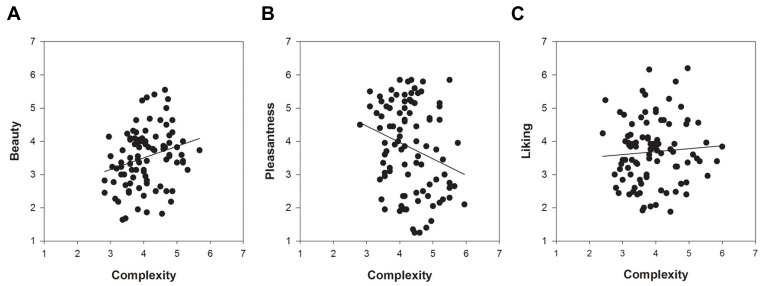

FIGURE 4 | The relationships between complexity and the three measures of hedonic tone for 94 representational paintings: (A) beauty, (B) pleasantness, and (C) liking.

$SD_{\text{pleasantness}} = 0.77$, $M_{\text{liking}} = 3.55$, $SD_{\text{liking}} = 0.74$; cartoons: $M_{\text{beauty}} = 3.76$, $SD_{\text{beauty}} = 0.61$, $M_{\text{pleasantness}} = 4.02$, $SD_{\text{pleasantness}} = 0.63$, $M_{\text{liking}} = 3.66$, $SD_{\text{liking}} = 0.61$; paintings: $M_{\text{beauty}} = 4.05$, $SD_{\text{beauty}} = 0.63$, $M_{\text{pleasantness}} = 4.31$, $SD_{\text{pleasantness}} = 0.68$, $M_{\text{liking}} = 3.90$, $SD_{\text{liking}} = 0.78$.

The analysis of the post-experimental questionnaire (**Table 5**) showed that participants' mean familiarity ratings were generally low (mean ratings around 2 across conditions on the 7-point rating scale). Furthermore, participants' liking ratings were neutral on average, which was to be expected given that pictures of different emotional contents were shown. In all conditions, participants reported to have felt feelings (mean ratings around 5 across conditions on the 7-point rating scale), which were of an average intensity. Moreover, participants also reported moderate emotional involvement with the picture content in all conditions. The difficulty to judge complexity was neutral across conditions. Ratings of the artistic status of cartoons showed that participants rated them above average (around 4.2 on a 7-point rating scale). A comparison across the three picture types by means of a series of Kruskal–Wallis tests revealed that participants neither preferred one picture type to the other nor did they report different levels of emotional involvement. All picture types induced emotions of similar intensity. However, the data suggest that reporting complexity judgments was somewhat easier for

TABLE 5 | Post-experiment questionnaire.

Group		General Liking	General Familiarity	Frequency of feelings	Intensity of feelings	Empathy	Difficulty to judge complexity	Artistic quality of cartoons
IAPS Beauty	*M*	4.17	1.83	5.33	4.50	5.50	4.63	–
(*n* = 24)	*SD*	1.37	1.17	1.24	1.14	1.14	1.38	
IAPS Pleasantness	*M*	4.71	2.67	5.52	5.05	5.86	4.95	–
(*n* = 21)	*SD*	1.31	1.39	0.75	0.81	0.73	1.60	
IAPS Liking	*M*	4.35	2.70	4.74	4.22	5.39	4.91	–
(*n* = 23)	*SD*	1.53	1.33	0.96	1.41	1.27	1.44	
Cartoons Beauty	*M*	3.88	2.48	4.76	3.88	5.28	4.44	4.20
(*n* = 25)	*SD*	1.59	1.39	1.42	1.24	1.37	1.71	1.66
Cartoons Pleasantness	*M*	4.18	2.73	5.00	4.41	5.41	4.55	4.27
(*n* = 22)	*SD*	1.56	1.98	1.02	1.37	0.80	1.47	1.80
Cartoons Liking	*M*	4.04	3.83	4.71	4.13	5.17	4.92	4.21
(*n* = 24)	*SD*	1.43	1.86	1.23	1.57	1.20	1.59	1.64
Paintings Beauty	*M*	3.95	2.05	4.5	3.91	5.14	3.73	–
(*n* = 22)	*SD*	1.36	1.00	1.26	1.80	1.49	1.39	
Paintings Pleasantness	*M*	4.75	2.35	5.25	4.60	5.80	4.00	–
(*n* = 20)	*SD*	1.56	1.31	1.07	1.05	0.77	1.32	
Paintings Liking	*M*	4.32	2.08	4.92	4.00	5.40	4.84	–
(*n* = 25)	*SD*	1.31	0.95	1.38	1.56	1.23	1.72	

Summary of the nine groups' mean answers to questions referring to general aspects of the experiment. M, mean; SD, standard deviation; n, number of participants; General liking: "How much did you like the pictures in general?"; General familiarity: "Please estimate how many pictures you have seen prior to the experiment."; Frequency of feelings: "Did you experience any feelings while looking at the pictures?"; Intensity of feelings: "How strong were the feelings evoked by the pictures?"; Empathy: "Did you feel emotionally involved with the visual contents of the pictures?"; Difficulty to judge complexity: "How difficult was it to rate the complexity of the pictures?"; Artistic quality of cartoons: "Would you consider these pictures as art?"

paintings than for the other two picture types, $H(2) = 5.70$, $p = 0.058$. General familiarity ratings were higher for cartoons than for the other two picture sets, $H(2) = 6.35$, $p = 0.042$, which was due to higher familiarity ratings in the condition in which participants reported liking of cartoons.

Discussion

We studied the relationship between complexity and three measures of hedonic tone in three sets of visual stimuli, namely environmental scenes, environmental scenes converted into cartoons, as well as representational paintings. In general, the current findings corroborated our hypotheses: After partialling out familiarity effects, we showed that visual complexity related to the number and diversity of elements and arousal were related in all three picture sets, and further, that this relationship was strongest for representational paintings, reaching a moderate effect size. The data also revealed that the multifaceted nature of hedonic tone was only clearly present during the perception of paintings: although the relationship between complexity and arousal was the same in all three conditions of hedonic tone, the relationship between complexity and hedonic tone differed across conditions. In line with Nadal et al. (2010), we observed a positive relationship between complexity and beauty. As predicted based on results by Marin and Leder (2016), we did not detect a significant relationship between complexity and pleasantness, although there were indications of a negative association between these measures if familiarity effects were ignored. Complexity and liking did not correlate with each other.

Further support for the multifaceted nature of hedonic tone in relation to paintings stems from the fact that the strength of the relationship between arousal and the three measures of hedonic tone differed considerably. The negative correlation was strongest between pleasantness and arousal, followed by the one between liking and arousal, and interestingly, there was no relationship between beauty and arousal. The relationship between the specific measure of hedonic tone and complexity during the experience of affective, representational human artworks may thus relate to hedonic measures' differential effects on arousal, a core component of Berlyne's theory. Moreover, the strength of the positive intercorrelations between the three measures of hedonic tone was lowest in paintings, although it was generally still high.

Marin and Leder (2016) investigated subjective responses to the same set of affective environmental scenes and representational paintings as used in the current study. Since the presentation duration of the pictures (5 s), the participant sample, as well as one measure of hedonic tone, namely pleasantness, were similar in the two studies, a detailed comparison of results is meaningful. The current results closely replicate findings by Marin and Leder (2016), who reported a correlation of $r_s = 0.27$ for the relationship between complexity and arousal in IAPS pictures and a correlation of $r_s = 0.52$ for the same relationship

in paintings. Moreover, the current findings are in line with the weakly negative relationship ($r_s = -0.15$) between complexity and pleasantness observed for paintings. However, Marin and Leder (2016) reported a weak positive relationship between complexity and pleasantness for IAPS pictures ($r_s = 0.22$), which was not supported by the current data, and also not by Marin and Leder (2013). It should be noted that Marin and Leder (2016) did not find a relationship between complexity and pleasantness for the presentation durations of 1 and 25 s, so we regard this finding as a possible random outcome. Altogether, the comparison of results across the two studies indicates that the nature of the relationship between complexity and arousal seems to be robust for these affective stimuli sets in groups of female participants.

Our findings are in line with previous studies employing more than one measure of hedonic tone, suggesting that people differentiate more between concepts such as beauty, pleasantness and liking during the perception of paintings (Berlyne and Ogilvie, 1974; Russell and George, 1990) than during the perception of real-life scenes or abstract visual patterns (Normore, 1974; Imamoglu, 2000). Here, we extend these findings by employing a strictly controlled research design that directly compares subjective responses across different stimulus sets within the same experimental framework. However, it still remains to be studied whether the multifaceted nature of hedonic tone and its relation to complexity is specific to esthetic experiences of human made artworks. The current data suggest that ratings of beauty, pleasantness and liking are nearly identical in the perception of affective and motivationally relevant environmental scenes, which probably have not evoked esthetic experiences. Therefore, our results motivate the systematic, comparative study of other stimulus categories in the visual domain.

The current results regarding the comparison across different picture sets are relevant to neuroimaging studies reporting brain areas being specifically active during the perception of art. Several neuroimaging studies have recently followed a comparative approach and investigated brain regions active during the perception of art and non-art images matched for semantic contents (Di Dio et al., 2011; Lacey et al., 2011; Lutz et al., 2013; Mizokami et al., 2014). Di Dio et al. (2011) investigated neural responses to masterpieces of classical masters and matched photographs of young athletes. In general, the activation patterns for the two types of stimuli were very similar; however, only artistic stimuli activated the right dorsal anterior insula, a brain area that has been found to link emotion and cognition. In a similar vein, Lutz et al. (2013) examined body representations in paintings and matched non-artistic photographs of body parts. The results indicated that processing of paintings was accompanied by the activation of the right parietal cortex and the extrastriate cortex bilaterally, which led the authors to conclude that the experience of visual art is a distinct perceptual process. Lacey et al. (2011) carefully matched photographs of everyday life with paintings and asked for several types of ratings (beauty, liking and pleasantness) after the fMRI scanning. The authors found activation of the ventral striatum for paintings but not for photographs. Interestingly, the authors did not report activation of the amygdala, and further, did not report correlations between the brain activity patterns and subjective ratings of beauty, pleasantness and liking. This suggests that activation of the ventral striatum may be related to the status of images as art and not to the esthetic experience itself as measured by subjective ratings. Mizokami et al. (2014) created visual scenes that very closely representing the semantic content of the landscapes and still life's and reported no correlation between beauty ratings and brain activations. Contrasting activations for paintings and photographs showed activations of the bilateral cuneus and the left lingual gyrus.

These previous studies comparing the neural correlates of different visual stimulus categories have largely ignored the affective contents of the stimuli and mostly focused on the matching of the semantic contents between categories. If one decided to follow an affective approach to the study of visual art, it would be crucial to incorporate pleasantness and arousal in the research design and stimulus selection, given that distinct neural networks related to the processing of pleasant and unpleasant environmental scenes (IAPS pictures) have been described (Aldhafeeri et al., 2012). The processing of pleasant pictures yielded significant activation in the bilateral prefrontal cortex, including the superior, medial and middle frontal gyri. Other active brain regions comprised the right anterior and posterior cingulate gyri and both temporal lobes. Unpleasant pictures elicited bilateral activation in the amygdala, hippocampus, parahippocampal gyri as well as secondary and primary visual cortex. These neural networks could be taken as a starting point for a neuroimaging study employing representational paintings. We further propose that future neuroimaging studies comparing different stimulus categories may also consider structural features of visual stimuli, such as aspects of complexity, and their relation to measures of hedonic tone. We would hypothesize that distinct activation patterns may emerge in the perception of paintings in comparison to the perception of matched non-art images. Such an approach would also contribute to a better understanding of the complex interplay between cognition and emotion during esthetic experiences.

EXPERIMENT 2

Materials and Methods

Participants

Participants were 92 female psychology students of the University of Vienna, all non-musicians (less than 3 years of musical training in the past and no musical activity at the time of the experiment). Participants were randomly assigned to one of three groups differing in the type of hedonic tone to be rated (beauty, pleasantness and liking). Participants were screened for outliers with respect to their mood prior to the experiment (participants with very low mood scores were removed), their musical sophistication as well as the trait emotional intelligence, emotional self-efficacy and SRSs. Fifteen participants were removed for these reasons by investigating the respective boxplots. The remaining 77 participants in the three groups did not significantly differ regarding their age, mood prior

TABLE 6 | Participant characteristics.

Group	Age [years]		Pos./neg. mood A	Alertness/ fatigue A	Quietude/ disquietude A	Gold-MSI	Trait EI	ES	SRS
Beauty	*M*	21.0	16.31	12.54	15.73	56.00	144.27	86.00	62.38
(*n* = 26)	*SD*	1.8	2.24	2.61	2.82	10.80	19.18	10.11	9.83
Pleasantness	*M*	20.5	16.72	13.64	15.20	53.24	146.60	86.04	60.92
(*n* = 25)	*SD*	1.4	2.28	2.87	2.75	14.58	15.14	8.52	7.76
Liking	*M*	21.8	17.54	12.62	15.00	53.88	146.96	84.23	61.27
(*n* = 26)	*SD*	2.8	1.68	3.29	3.03	15.25	17.51	10.33	6.26
Test statistic	*H* = 1.85		*H* = 4.30 2	*H* = 1.89	*H* = 1.02	*F* = 2.89	*F* = 0.18	*F* = 0.29	*F* = 0.23
p	0.396		0.116	0.389	0.601	0.752	0.833	0.747	2,74
df	2		2	2	2	2,74	2,74	2,74	2,74

Summary of the three groups' mean age, mood prior to the experiment (MDBF short form A), musical sophistication, trait emotional intelligence, emotional self-efficacy and stress reactivity. M, mean; SD, standard deviation; n, number of participants; p, calculated probability; df, degrees of freedom; H, Kruskal–Wallis test statistic; F, F-test statistic; Gold-MSI, Goldsmiths Musical Sophistication Index; pos., positive; neg., negative; EI, emotional intelligence; ES, emotional self-efficacy; SRS, stress reactivity scale.

TABLE 7 | Intra-class correlation coefficients (ICCs) for subjective ratings of familiarity, complexity, arousal and hedonic tone (beauty, pleasantness and liking).

Group	ICC(2,k) Familiarity	ICC(2,k) Complexity	ICC(2,k) Arousal	ICC(2,k) Hedonic tone
Beauty	0.717	0.897	0.923	0.810
(*n* = 26)	CI [0.626,0.794]	[0.864,0.925]	[0.898,0.944]	[0.749,0.862]
Pleasantness	0.742	0.934	0.926	0.851
(*n* = 25)	CI [0.659,0.812]	[0.913,0.952]	[0.903,0.946]	[0.804,0.892]
Liking	0.776	0.952	0.935	0.820
(*n* = 26)	CI [0.705,0.837]	[0.936,0.965]	[0.915,0.953]	[0.763,0.869]

ICC(2,k), two-way random average measure, type consistency; CI, 95% confidence interval. ICCs were calculated for ratings of 92 musical excerpts.

to the experiment and their scores on the trait EI, emotional self-efficacy and SRS. We also ensured that musical sophistication was similar and generally low (see **Table 6**).

Materials

Ninety-two musical excerpts, taken from Experiment 3 of Marin and Leder (2013), were shortened to a duration of 15 s by removing the endings of the original excerpts. A fade-out (500 ms) was added to the excerpts using Audacity 2.0.6 software. These musical stimuli varied naturally in complexity and were pre-selected to cover the two-dimensional emotion space spanned by pleasantness and arousal (Russell, 1980).

As in Experiment 1, several standardized self-report measures were administered, including the two short forms A and B of the multidimensional mood questionnaire (MDBF, Steyer et al., 1997), the short trait emotional intelligence questionnaire (TEIQue-SF, Petrides and Furnham, 2006), the stress reactivity scale (SRS, Schulz et al., 2005), and the emotional self-efficacy scale by Schmitz and von Salisch (2002). Musical sophistication was assessed with the German version of the Goldsmiths Musical Sophistication Index (Gold-MSI) developed by Schaal et al. (2014).

A self-developed questionnaire comprised a set of questions to be answered on 7-point scales referring to the general liking of the musical excerpts, the difficult of judging musical complexity and arousal, and the strengths of the feelings induced by the music. Participants were also asked to report on the role of music in their lives and preferences for different musical styles.

Procedure

As in Experiment 1, we followed the current ethical guidelines at the University of Vienna and the version of the Ethical Principles of Psychologists and Code of Conduct of the American Psychological Association (with 2010 amendments). All participants read and signed an informed consent form.

Participants filled in the short form A of the multidimensional mood questionnaire prior to the experiment. The experimental setting was identical to the one described in Experiment 1 except that musical stimuli were played through an external soundcard (E-MU audio interface, 0204/USB) and participants were wearing Sennheißer HD 380 pro headphones. The volume was fixed to approximately 72 dB SPL (A-weighted) as measured during the presentation of the second practice trial (see Marin and Leder, 2013). Sitting 60–70 cm away from the screen (19-inch Iiyama ProLite B1906S), participants were familiarized with the experimental task in two practice trials. Each trial was announced for 5 s and then the excerpt was played for

15 s, after which the first (out of four) 7-point rating scale appeared on the screen (familiarity, complexity, arousal and hedonic tone). The instructions were identical as those used in Experiment 1. Familiarity ratings referred to the familiarity with the musical excerpt. All ratings were given by mouse click and self-paced.

The musical excerpts were blocked according to emotional contents (low-arousing pleasant, low-arousing unpleasant, high-arousing pleasant and high-arousing unpleasant). The order of the four blocks as well as the order of the 23 excerpts within each block were randomized. The participants were told that the blocks would vary in their emotional content, and they were allowed to take breaks between blocks. After the experiment, participants filled in the short form B of the mood questionnaire, and then the other questionnaires in one of two different orders (TEIQue-SF, SRS, emotional self-efficacy, and Gold-MSI or emotional self-efficacy, SRS, TEIQue-SF, and Gold-MSI). The experimental session lasted around 90 min, after which participants were debriefed, thanked and dismissed.

Results

The analysis of the data follows the one presented in Experiment 1. The subjective ratings were averaged across participants for each of the musical excerpt in each condition. Then we assessed the inter-rater reliability by calculating the average intra-class correlation coefficient (ICC) using a two-factor random effects model and type consistency (Shrout and Fleiss, 1979; McGraw and Wong, 1996). **Table 7** shows that the ICCs were generally high (all ICCs > 0.7), which justifies averaging across participants and considering excerpts as unit of analysis. Next, outliers were determined by examining boxplots for each type of rating for each condition. There were no outliers regarding familiarity, complexity and arousal ratings. One excerpt was removed due to very low pleasantness ratings, four excerpts due to very low beauty ratings, and one due to very low liking ratings. Since most distributions were not normal, non-parametric analyses were employed to investigate the interrelationships between the four types of ratings in each condition.

Table 8 depicts the interrelationships between the four types of ratings, separately given for each of the three conditions. All relationships were visually inspected to ensure that they were monotonous before Spearman's rank correlations were computed. Familiarity was only mildly negatively associated with complexity and arousal, whereas the relationships between familiarity and measures of hedonic tone were positive, ranging from medium (r_s ~0.53) to strong effect sizes (r_s ~0.75). The relationship between complexity and arousal was strong and of equal strength in all three conditions of hedonic tone (r_s ~0.85). The relationship between complexity and beauty was not linear (**Figure 5**). The relationship followed an inverted-U curve and curve-fitting analysis confirmed its quadratic nature, $F(2,85) = 5.24$, $p = 0.007$, $R^2 = 0.11$, $y = 0.9 + 1.58x - 0.18x^2$. Complexity and pleasantness were significantly negatively associated (r_s ~−0.25), and the data revealed no relationship between complexity and liking.

Controlling for familiarity effects, **Table 9** shows how complexity, arousal and the three measures of hedonic tone were associated. The strong and positive association between complexity and arousal is in line with Berlyne's theory. Three different relationships between complexity and the respective measures of hedonic tone emerged: complexity and pleasantness were negatively associated (r_s ~−0.26), complexity and liking positively (r_s ~0.29), and finally, complexity and beauty ratings followed an inverted-U curve as predicted by Berlyne. The two linear relationships significantly differed from each other, $z = -3.72$, $p = 0.0002$. Next, a hierarchical linear regression analysis was performed to control for effects of familiarity on beauty ratings and to demonstrate the quadratic nature of the relationship. Familiarity was entered in the first block, followed by adding the linear term for complexity ratings, and finally, the quadratic term for complexity was entered in the third step. The results of the final model were significant, $F(3,87) = 17.31$, $p < 0.001$, adjusted $R^2 = 0.36$, and revealed the following significant predictors: familiarity ($\beta = 0.54$, $p < 0.001$), complexity ($\beta = 1.72$, $p = 0.041$) and complexity squared ($\beta = -2.12$, $p = 0.037$).

Table 9 also shows that the relationships between arousal and measures of hedonic tone differed from each other. The relationship between arousal and beauty was of a non-monotonic quadratic nature, though curve-fitting analysis did not reveal a significant trend. The relationship between pleasantness and arousal was negative (r_s ~−0.44), and the one between arousal and liking positive (r_s ~0.24). In general, the correlations between arousal and measures of hedonic tone reflect the nature of the complexity-hedonic tone relationship.

Next, intercorrelations between measures of hedonic tone were computed ($N = 88$), with the lowest correlation between liking and pleasantness ratings: $r_{\text{beauty-pleasantness}} = 0.70$, $r_{\text{beauty-liking}} = 0.76$, and $r_{\text{liking-pleasantness}} = 0.59$. Intercorrelations between the three complexity ratings of the three conditions were generally very high: $r_{\text{beauty-pleasantness}} = 0.96$, $r_{\text{beauty-liking}} = 0.96$, and $r_{\text{liking-pleasantness}} = 0.96$. For arousal ratings, these intercorrelations across the three conditions were of similar strength: $r_{\text{beauty-pleasantness}} = 0.94$, $r_{\text{beauty-liking}} = 0.95$, and $r_{\text{liking-pleasantness}} = 0.95$. The mean complexity ratings of the three conditions were as follows: $M_{\text{beauty}} = 4.35$, $SD_{\text{beauty}} = 0.88$,

TABLE 8 | Spearman's rank order correlations between measures of familiarity, complexity, arousal and measures of hedonic tone in response to 19th-century piano solo music.

Group	*df*	Measure	Familiarity	Complexity
Beauty	86	Complexity	0.10	
Pleasantness	89		−0.08	
Liking	89		−0.21	
Beauty	86	Hedonic tone	0.53*	N/A
Pleasantness	89		0.64*	−0.25*
Liking	89		0.75*	0.04
Beauty	86	Arousal	0.003	0.86*
Pleasantness	89		−0.30*	0.85*
Liking	89		−0.35*	0.85*

$p < 0.05$ after Bonferroni–Holm correction; df, degrees of freedom.

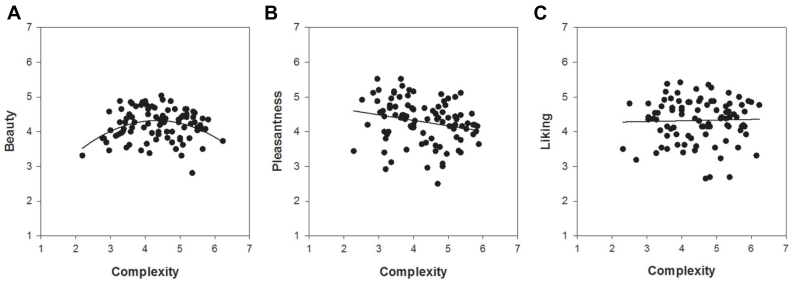

FIGURE 5 | The relationships between complexity and the three measures of hedonic tone for 19th-century piano solo music: (A) beauty, (B) pleasantness, and (C) liking.

$M_{\text{pleasantness}} = 4.27$, $SD_{\text{pleasantness}} = 0.91$, $M_{\text{liking}} = 4.49$, $SD_{\text{liking}} = 0.94$.

Participants also answered several questions referring to the experiment (**Table 10**). The three groups did not differ regarding their general liking of the musical excerpts, their familiarity with them, the frequency and strength of reported feelings as well as the difficulty of judging musical complexity. Significant group differences were observed for the role of music in participants' life, $H(2) = 8.64$, $p = 0.013$, as well as for the frequency of listening to classical music, $H(2) = 7.73$, $p = 0.021$. Participants in the beauty group reported a higher role of music in their lives and a lower frequency of listening to music in comparison to the other two groups.

Discussion

We explored the relationship between musical complexity and three measures of hedonic tone (beauty, pleasantness and liking) in a highly controlled participant sample and between-subject design. The same stimulus set was rated for familiarity, complexity, arousal and one of the three measures of hedonic tone. We observed that the nature of the relationship differed across groups when controlling for familiarity effects. Not supporting our hypotheses, but in line with Berlyne's theory (Berlyne, 1971), our data indicated an inverted-U relationship between complexity and beauty ratings. Moreover, we observed a weak negative relationship between complexity and pleasantness ratings and a weak positive relationship between complexity and liking. In other words, all three measures of hedonic tone showed different relationships with complexity. Arousal and complexity were highly correlated, as predicted by Berlyne, and the effect size was not only nearly identical across the three groups, but also very similar to results reported in Experiment 3 in Marin and Leder (2013).

TABLE 9 | Partial Spearman's rank order correlations controlling for familiarity between measures of complexity, arousal and hedonic tone in response to 19th-century piano solo music.

Group	*df*	Measure	Complexity	Hedonic tone
Beauty	85	Hedonic tone	N/A	
Pleasantness	88		−0.26*	
Liking	88		0.29**	
Beauty	85	Arousal	0.87***	N/A
Pleasantness	88		0.87***	−0.44***
Liking	88		0.85***	0.24*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; *df*, degrees of freedom.

Marin and Leder (2013) studied the relationship between complexity and pleasantness using longer versions of the musical excerpts employed in Experiment 2. After controlling for familiarity effects, their results indicated no relationship between complexity and pleasantness in females, but a positive relationship between these measures in males. Here, we found a negative relationship between complexity and pleasantness. Because we employed a more rigorous approach during participant sampling than in Marin and Leder (2013), it may be that other factors such as mood or musical sophistication influenced previously reported results.

The present results further show that the relationship between arousal and measures of hedonic tone is comparable to the respective relationships between complexity and hedonic tone (see **Table 9**), lending more support to Berlyne's theory regarding the crucial role of arousal in esthetic experiences. This finding was also reported in Marin and Leder (2013) and thus clearly suggests that arousal plays a key role in the complexity-hedonic tone relationship. Future research may investigate larger samples of participants in order to test whether arousal acts as a mediator in the complexity-hedonic tone relationships. Using stimuli as the unit of analysis, Marin and Leder (2013) already provided some support for Berlyne's theory by testing mediation with the help of regression analysis. However, it may be more appropriate to test mediation using individual participant data, which requires a much larger sample than the one tested in this study.

The stimulus selection in Marin and Leder (2013) primarily focused on the emotional content of the stimuli, following Russell's (1980) circumplex model of affect, and no particular attention was paid to the complexity of the musical excerpts. This stands in contrast to the approach presented in Experiment 1, where complexity was considered during the stimulus selection and differed mostly along the dimension of elements (Marin and Leder, 2016). Since it is not known yet whether the perception of musical complexity can be understood using

TABLE 10 | Post-experiment questionnaire.

Group	General Liking		General Familiarity	Frequency of feelings	Intensity of feelings	Difficulty to judge complexity	Role of music in life	Listening to classical music
Beauty	*M*	4.46	2.34	4.77	3.88	4.65	5.88	2.15
(*n* = 26)	*SD*	1.45	0.80	1.39	1.48	1.62	0.86	1.08
Pleasantness	*M*	4.88	3.08	4.72	3.68	5.04	5.12	2.76
(*n* = 25)	*SD*	1.09	1.08	1.14	0.95	1.37	1.17	1.30
Liking	*M*	5.08	2.77	4.50	3.69	4.23	5.00	3.23
(*n* = 26)	*SD*	1.55	1.50	1.50	1.59	1.86	1.26	1.48

Summary of the three groups' mean answers to questions referring to general aspects of the experiment. M, mean; SD, standard deviation; n, number of participants; General liking: "How much did you like the music in general?"; General familiarity: "Please estimate how many musical excerpts you have heard prior to the experiment."; Frequency of feelings: "Did you experience any feelings while listening to the music?"; Intensity of feelings: "How strong were the feelings evoked by the music?"; Difficulty to judge complexity: "How difficult was it to rate the complexity of the musical excerpts?", Role of music in life: "What role does music play in your life?"

dimensions comparable to those in the visual domain, such as number and diversity of elements or disorganization, a direct comparison between the specific outcomes for each of the hedonic measures across the visual and musical domains must be made with caution. However, our findings strongly suggest that in the appreciation of paintings and music, the nature of the complexity-hedonic tone relationship depends on the specific measure of hedonic tone if complexity is held constant.

GENERAL DISCUSSION

We sought to contribute to a better understanding of the discrepant research results regarding Berlyne's (1971) psychobiological model of esthetic experience. Thus, following a comparative approach (Marin, 2015), we systematically examined the nature of hedonic tone in the experience of affective environmental scenes, environmental scenes converted to cartoons, representational paintings as well as music. One single factor of esthetic preference has previously been proposed, and beauty has been described as the concept best representing this factor (Eysenck, 1940; Marty et al., 2003; Jacobsen et al., 2004; Augustin et al., 2012a,b). Empirical evidence for this claim was usually based on results of a factor analysis, which showed that different evaluative scales loaded on one common underlying factor, and that beauty had the strongest loading on this factor. Consequently, researchers have tended to consider mainly beauty as a measure of hedonic tone in the study of esthetic experiences (Nadal et al., 2010). In the present study, we explored the possibility that hedonic tone has a multifaceted nature that cannot be ignored in modeling esthetic experiences.

Our primary hypothesis was that beauty and other measures of hedonic tone might not be similarly related to complexity if the latter is held constant. In a between-subjects design, subjective ratings of beauty, pleasantness and liking were collected for each stimulus set and correlated with ratings of complexity by controlling for familiarity effects. We predicted that the multifaceted nature of hedonic tone might be more evident in responses to paintings, cartoons and music than in environmental scenes depicting common, everyday life scenes. The different aspects of hedonic tone only became clearly apparent with regard to human-made visual artworks and music, for which we observed different relationships with complexity, ranging from positive and negative linear relationships, to quadratic relationships as well as indications of no relationship between these variables. Even subjective responses to cartoon-like pictures based on environmental scenes did not show a significant effect (although we observed weak indications), possibly due to the fact that cartoons were not considered as having a high esthetic quality or artistic status by participants.

With respect to Berlyne's (1971) psychobiological model, our data suggests that the proposed inverted-U curve between complexity and hedonic tone may only be one out of several possible associations, depending on the measure of hedonic tone as well as the underlying dimension of complexity. However, the positive linear relationship between complexity and arousal was confirmed by our data across several stimulus categories and participants groups. As suggested by Nadal et al. (2010), sub-dimensions of visual complexity may show different associations with beauty ratings of visual stimuli. Here, we further demonstrate that the relationship with complexity (if held constant) may differ across measures of hedonic tone. Taken together, these results indicate that a differentiated view of concepts such as complexity and hedonic tone is warranted. It is plausible to assume that, for instance, dimensions of visual complexity (Nadal et al., 2010) show different relationships with beauty, liking and pleasantness. Likewise, the experience of musical complexity may by generated by several underlying dimensions whose associations with measures of hedonic tone may differ.

It is difficult to speculate why hedonic measures are not similarly associated with complexity, although they are themselves highly correlated in all experimental conditions and across stimulus sets. Esthetic experiences are currently being understood as interplay between emotion and cognition as well as interplay between bottom–up and top–down processes. For example, it is possible that beauty, pleasantness and liking differ in their affective connotations, with pleasantness presumably most closely linked to emotion. It can be surmised that the relationship between complexity and cognitive processing is partly determined by the affective content of the hedonic tone measures. Another related explanation refers to the different

types of associations we observed between arousal and measures of hedonic tone in both paintings and music. If Berlyne was right to assume that arousal has a key function in esthetic experiences, then the nature of the arousal-hedonic tone relationship may play a role in the relationship between complexity and hedonic tone. Furthermore, it remains to be seen whether measures of hedonic tone contribute to one type of subjective experience in the perception of visual arts and music, which could be labeled as hedonic tone, or whether these concepts contribute to distinct experiences that are somewhat correlated and part of an esthetic experience, which may comprise other aspects than hedonic tone. This approach would also lead to a better understanding of whether and how visually and musically induced esthetic experiences differ.

To the best of our knowledge, no neuroimaging study has compared the activation patterns related to different hedonic concepts and complexity for the same picture set, let alone examined functional connectivity patterns. Our current results thus allow for the formulation of concrete hypotheses for future neuroimaging studies. For instance, since we present evidence that hedonic tone has a multifaceted nature in the experience of visual art and music, the respective underlying neural substrates may slightly differ and perhaps be dissimilarly connected to other brain regions involved in esthetic experiences, particularly those involved in structural feature processing. To be specific, paradigms used by Jacobsen and Höfel (2003) and Jacobsen et al. (2006) who compared the neural correlates of beauty judgments with those of symmetry judgments, may be extended by incorporating other measures of hedonic tone as well as by considering the multidimensionality of complexity (Nadal et al., 2010). We speculate that brain activations related to pleasantness may very likely differ from those related to beauty because the former can be considered as a dimension of core affect (Russell, 2009). This is also reflected in our data because we found a stronger correlation between pleasantness and arousal than between beauty and arousal. Further support for this argument comes from related research in the music domain. Brattico et al. (2013) developed a neurobiological model of esthetic experiences in music and argue that core affect is processed in brain areas such as the amygdala, sensory cortices and the parahippocampal gyrus, whereas esthetic judgments of beauty may primarily activate the orbitofrontal cortex, the anterior cingulate and the premotor cortex. Their model further proposes that brain areas active in beauty judgments may be active during reports of liking occurring during later processing stages. Reports of liking are also associated with activations in the ventral striatum and the insula. In order to fully understand this co-activation of brain areas involved in the processing of beauty and liking a systematic study of the time course of these processes will be crucial.

The role of arousal during esthetic processing has largely been ignored in neuroaesthetics, not only in the musical domain, as pointed out by Brattico et al. (2013), but also in the study of visual esthetic experiences (Leder et al., 2015), in which much emphasis has been placed on different aspects of valence, such as "positive vs. negative," "happy vs. sad" or "pleasant vs. unpleasant." In other words, a systematic study of the contributing affect systems during esthetic experiences that considers not only valence but also arousal is warranted (Russell, 1980). It may be true that in the context of Berlyne's theory, which has arousal at its core, the relationship between hedonic tone and complexity may not always follow an inverted-U curve; however, arousal may still play a decisive role during esthetic episodes (Marin and Leder, 2013). Consequently, we highlight not only the need to study autonomic arousal but also the investigation of brain correlates of subjective arousal levels that may be associated with different neural affect systems (Koelsch et al., 2015) during esthetic experiences.

The present study also sheds more light on the role of familiarity in esthetic experiences and clearly indicates that it is necessary to account for familiarity effects even though participants may generally be unfamiliar with the stimuli, and as it was the case here, were non-experts in visual arts and non-musicians. In both the visual and musical domains, we observed on average moderate positive correlations between familiarity and measures of hedonic tone, and small negative correlations between familiarity and arousal as well as familiarity and complexity. Consequently, it is crucial to account for this effect by statistical techniques if one is interested in the nature of the complexity-hedonic tone relationship across stimulus categories and different participant groups.

We are aware of several limitations of the current study impeding broader generalizations. In both experiments, as well as in our previous work (Marin and Leder, 2013, 2016), we followed an affective approach by selecting stimuli varying in their affective contents (i.e., in arousal and pleasantness). It cannot be ruled out that our results are specific to these types of stimuli, and that findings for neutral stimuli may differ in the visual domain. Future studies may also involve male subjects and those showing high art interest or musical sophistication, other stimuli types, such as abstract artworks, and other measures of hedonic value, such as reward value. Furthermore, psychophysiological and neurophysiological measures may be added to the research design, especially since Berlyne (1974a) already proposed different psychophysiologcial signatures of measures of hedonic tone.

To conclude, this systematic study demonstrates that discrepancies regarding Berlyne's psychobiological model might partly be rooted in the largely ignored multifaceted nature of hedonic tone. Future research, especially in the growing field of neuroaesthetics, may thus explore these nuanced aspects in greater depth.

AUTHOR CONTRIBUTIONS

MM designed the experiment and created the stimulus library. AL, MW, and MM collected the data. MM analyzed and interpreted the data. MM drafted the work and revised it after feedback from AL, MW, and HL. All authors provided final approval and agree to be accountable for all aspects of the work.

FUNDING

This article was supported by the Open Access Publishing Fund of the University of Vienna.

ACKNOWLEDGMENTS

We wish to thank Balint Puster for his assistance in data collection and Bruno Gingras for useful comments on an earlier version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum.2016.00536/full#supplementary-material

REFERENCES

Aldhafeeri, F. M., Mackenzie, I., Kay, T., Alghamdi, J., and Sluming, V. (2012). Regional brain responses to pleasant and unpleasant IAPS pictures: different networks. *Neurosci. Lett.* 512, 94–98. doi: 10.1016/j.neulet.2012.01.064

Augustin, M. D., Carbon, C.-C., and Wagemans, J. (2012a). Artful terms: a study on aesthetic word usage for visual art versus film and music. *i-Perception* 3, 319–337. doi: 10.1068/i0511aap

Augustin, M. D., Wagemans, J., and Carbon, C.-C. (2012b). All is beautiful? Generality vs. specificity of word usage in visual aesthetics. *Acta Psychol.* 139, 187–201. doi: 10.1016/j.actpsy.2011.10.004

Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York, NY: Appleton-Century-Crofts.

Berlyne, D. E. (1974a). *Studies in the New Experimental Aesthetics: Steps Toward an Objective Psychology of Aesthetic Appreciation*. Washington, DC: Hemisphere.

Berlyne, D. E. (1974b). "Verbal and exploratory responses to visual patterns varying in uncertainty and in redundancy," in *Studies in the New Experimental Aesthetics*, ed. D. E. Berlyne (New York, NY: John Wiley & Sons), 121–158.

Berlyne, D. E., and Ogilvie, J. C. (1974). "Dimensions of perception of paintings," in *Studies in the New Experimental Aesthetics*, ed. D. E. Berlyne (New York, NY: John Wiley & Sons), 181–226.

Bertamini, M., Makin, A., and Rampone, G. (2013). Implicit association of symmetry with positive valence, high arousal and simplicity. *i-Perception* 4, 317–327. doi: 10.1068/i0601jw

Blijlevens, J., Carbon, C.-C., Mugge, R., and Schoormans, J. P. L. (2012). Aesthetic appraisal of product designs: independent effects of typicality and arousal. *Br. J. Psychol.* 103, 44–57. doi: 10.1111/j.2044-8295.2011.02038.x

Boccia, M., Barbetti, S., Piccardi, L., Guariglia, C., Ferlazzo, F., Giannini, A. M., et al. (2016). Where does brain neural activation in aesthetic responses to visual art occur? Meta-analytic evidence from neuroimaging studies. *Neurosci. Biobehav. Rev.* 60, 65–71. doi: 10.1016/j.neubiorev.2015.09.009

Boccia, M., Nemmi, F., Tizzani, E., Guariglia, C., Ferlazzo, F., Galati, G., et al. (2015). Do you like Arcimboldo's? Esthetic appreciation modulates brain activity in solving perceptual ambiguity. *Behav. Brain Res.* 278, 147–154. doi: 10.1016/j.bbr.2014.09.041

Brattico, E., Bogert, B., and Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Front. Psychol.* 4:206. doi: 10.3389/fpsyg.2013.00206

Cela-Conde, C. J., Agnati, L., Huston, J. P., Mora, F., and Nadal, M. (2011). The neural foundations of aesthetic appreciation. *Prog. Neurobiol.* 94, 39–48. doi: 10.1016/j.pneurobio.2011.03.003

Cela-Conde, C. J., and Ayala, F. J. (2014). Brain keys in the appreciation of beauty: a tale of two worlds. *Rend. Lincei* 25, 277–284. doi: 10.1007/s12210-014-0299-8

Cela-Conde, C. J., García-Prieto, J., Ramasco, J. J., Mirasso, C. R., Bajo, R., Munar, E., et al. (2013). Dynamics of brain networks in the aesthetic appreciation. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl. 2), 10454–10461. doi: 10.1073/pnas.1302855110

Cela-Conde, C. J., Marty, G., Maestu, F., Ortiz, T., Munar, E., Fernandez, A., et al. (2004). Activation of the prefrontal cortex in the human visual aesthetic perception. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6321–6325. doi: 10.1073/pnas.0401427101

Chatterjee, A., and Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn. Sci.* 18, 370–375. doi: 10.1016/j.tics.2014.03.003

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*, 2nd Edn. Hillsdale, NJ: Erlbaum.

Cupchik, G. C. (1988). The legacy of Daniel E. Berlyne. *Empir. Stud. Arts* 6, 171–186. doi: 10.2190/FLM8-6NQ7-N5WM-WLLT

Cupchik, G. C., Vartanian, O., Crawley, A., and Mikulis, D. J. (2009). Viewing artworks: contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain Cogn.* 70, 84–91. doi: 10.1016/j.bandc.2009.01.003

Di Dio, C., Canessa, N., Cappa, S. F., and Rizzolatti, G. (2011). Specificity of esthetic experience for artworks: an fMRI study. *Front. Hum. Neurosci.* 5:139. doi: 10.3389/fnhum.2011.00139

Eisenberg, J., and Thompson, W. F. (2003). A matter of taste: evaluating improvised music. *Creat. Res. J.* 15, 287–296. doi: 10.1080/10400419.2003.9651421

Eysenck, H. J. (1940). The general factor in aesthetic judgements. *Br. J. Psychol. Gen. Sect.* 31, 94–102. doi: 10.1111/j.2044-8295.1940.tb00977.x

Flexas, A., Rosselló, J., Christensen, J. F., Nadal, M., La Rosa, A. O., and Munar, E. (2013). Affective priming using facial expressions modulates liking for abstract art. *PLoS ONE* 8:e80154. doi: 10.1371/journal.pone.0080154

Gartus, A., and Leder, H. (2014). The white cube of the museum versus the gray cube of the street: the role of context in aesthetic evaluations. *Psychol. Aesthet. Creat. Arts* 8, 311–320. doi: 10.1037/a0036847

Gingras, B., Marin, M. M., and Fitch, W. T. (2014). Beyond intensity: spectral features effectively predict music-induced subjective arousal. *Q. J. Exp. Psychol.* 67, 1428–1446. doi: 10.1080/17470218.2013.863954

Graf, L. K. M., and Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking. *Pers. Soc. Psychol. Rev.* 19, 395–410. doi: 10.1177/1088868315574978

Güçlütürk, Y., Jacobs, R. H., and Van Lier, R. (2016). Liking versus complexity: decomposing the inverted U-curve. *Front. Hum. Neurosci.* 10:112. doi: 10.3389/fnhum.2016.00112

Hagtvedt, H., Patrick, V. M., and Hagtvedt, R. (2008). The perception and evaluation of visual art. *Empir. Stud. Arts* 26, 197–218. doi: 10.2190/EM.26.2.d

Heath, T., Smith, S. G., and Lim, B. (2000). Tall buildings and the urban skyline the effect of visual complexity on preferences. *Environ. Behav.* 32, 541–556. doi: 10.1177/00139160021972658

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Imamoglu, Ç (2000). Complexity, liking and familiarity: architecture and non-architecture Turkish students' assessments of traditional and modern house facades. *J. Environ. Psychol.* 20, 5–16. doi: 10.1006/jevp.1999.0155

Ishizu, T., and Zeki, S. (2011). Toward a brain-based theory of beauty. *PLoS ONE* 6:e21852. doi: 10.1371/journal.pone.0021852

Jacobs, K. W., and Hustmyer, F. E. Jr. (1974). Effects of four psychological primary colors on GSR, heart rate and respiration rate. *Percept. Mot. Skills* 38, 763–766. doi: 10.2466/pms.1974.38.3.763

Jacobsen, T. (2006). Bridging the arts and sciences: a framework for the psychology of aesthetics. *Leonardo* 39, 155–162. doi: 10.1162/leon.2006.39.2.155

Jacobsen, T., Buchta, K., Köhler, M., and Schröger, E. (2004). The primacy of beauty in judging aesthetics of objects. *Psychol. Rep.* 94, 1253–1260. doi: 10.2466/PR0.94.3.1253-1260

Jacobsen, T., and Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cogn. Affect. Behav. Neurosci.* 3, 289–299. doi: 10.3758/CABN.3.4.289

Jacobsen, T., Schubotz, R. I., Höfel, L., and Cramon, D. Y. V. (2006). Brain correlates of aesthetic judgment of beauty. *Neuroimage* 29, 276–285. doi: 10.1016/j.neuroimage.2005.07.010

Kawabata, H., and Zeki, S. (2004). Neural correlates of beauty. *J. Neurophysiol.* 91, 1699–1705. doi: 10.1152/jn.00696.2003

Kirk, U. (2008). The neural basis of object-context relationships on aesthetic judgment. *PLoS ONE* 3:e3754. doi: 10.1371/journal.pone.0003754

Kirk, U., Skov, M., Christensen, M. S., and Nygaard, N. (2009). Brain correlates of aesthetic expertise: a parametric fMRI study. *Brain Cogn.* 69, 306–315. doi: 10.1016/j.bandc.2008.08.004

Koelsch, S., Jacobs, A. M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C., et al. (2015). The quartet theory of human emotions: an integrative and neurofunctional model. *Phys. Life Rev.* 13, 1–27. doi: 10.1016/j.plrev.2015.03.001

Konecni, V. J. (1978). Daniel E. Berlyne: 1924-1976. *Am. J. Psychol.* 91, 133–137.

Konecni, V. J. (1996). Daniel E. Berlyne (1924–1976): two decades later. *Empir. Stud. Arts* 14, 129–142. doi: 10.2190/DNFJ-8230-WBYX-JJB5

Konecni, V. J., Crozier, J., and Doob, A. (1976). Anger and expression of aggression: effects on aesthetic preference. *Sci. Aesthet.* 1, 47–55.

Konecni, V. J., and Sargent-Pollock, D. (1977). Arousal, positive and negative affect, and preference for Renaissance and 20th-century paintings. *Mot. Emot.* 1, 75–93. doi: 10.1007/BF00997582

Kreitler, S., Zigler, E., and Kreitler, H. (1974). The complexity of complexity. *Hum. Dev.* 17, 54–73. doi: 10.1159/000271333

Krupinski, E., and Locher, P. (1988). Skin conductance and aesthetic evaluative responses to nonrepresentational works of art varying in symmetry. *Bull. Psychon. Soc.* 26, 355–358. doi: 10.3758/BF03337681

Lacey, S., Hagtvedt, H., Patrick, V. M., Anderson, A., Silla, R., Deshpande, G., et al. (2011). Art for reward's sake: visual art recruits the ventral striatum. *Neuroimage* 55, 420–433. doi: 10.1016/j.neuroimage.2010.11.027

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective Picture System (IAPS). Affective Ratings of Pictures and Instruction Manual.* Technical report a-7. Gainesville, FL: University of Florida.

Leder, H., Belke, B., Oeberst, A., and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *Br. J. Psychol.* 95, 489–508. doi: 10.1348/0007126042369811

Leder, H., Markey, P. S., and Pelowski, M. (2015). Aesthetic emotions to art–what they are and what makes them special. comment on "The quartet theory of human emotions: an integrative and neurofunctional model" by S. Koelsch et al. *Phys. Life Rev.* 13, 67–70. doi: 10.1016/j.plrev.2015.04.037

Leder, H., and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode – developments and challenges in empirical aesthetics. *Br. J. Psychol.* 105, 443–464. doi: 10.1111/bjop.12084

Lutz, A., Nassehi, A., Bao, Y., Pöppel, E., Sztrókay, A., Reiser, M., et al. (2013). Neurocognitive processing of body representations in artistic and photographic images. *Neuroimage* 66, 288–292. doi: 10.1016/j.neuroimage.2012.10.067

Marin, M. M. (2015). Crossing boundaries: toward a general model of neuroaesthetics. *Front. Hum. Neurosci.* 9:443. doi: 10.3389/fnhum.2015.00443

Marin, M. M., and Leder, H. (2013). Examining complexity across domains: relating subjective and objective measures of affective environmental scenes. *Paintings Music. PLoS ONE* 8:e72412. doi: 10.1371/journal.pone.0072412

Marin, M. M., and Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychol.* 163, 38–58. doi: 10.1016/j.actpsy.2015.10.002

Marković, S. (2012). Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *i-Perception* 3, 1–17. doi: 10.1068/i0450aap

Martindale, C., and Moore, K. (1988). Priming, prototypicality, and preference. *J. Exp. Psychol.* 14, 661–670.

Martindale, C., Moore, K., and Borkum, J. (1990). Aesthetic preference: anomalous findings for Berlyne's psychobiological theory. *Am. J. Psychol.* 103, 53–80. doi: 10.2307/1423259

Marty, G., Conde, C. J. C., Munar, E., Rosselló, J., Roca, M., and Escudero, J. T. (2003). Dimensiones factoriales de la experiencia estética. *Psicothema* 15, 478–483.

McGraw, K. O., and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* 1, 30–46. doi: 10.1037/1082-989X.1.1.30

Messinger, S. M. (1998). Pleasure and complexity: Berlyne revisited. *J. Psychol.* 132, 558–560. doi: 10.1080/00223989809599288

Mizokami, Y., Terao, T., Hatano, K., Hoaki, N., Kohno, K., Araki, Y., et al. (2014). Difference in brain activations during appreciating paintings and photographic analogs. *Front. Hum. Neurosci.* 8:478. doi: 10.3389/fnhum.2014.00478

Munar, E., Nadal, M., Rossello, J., Flexas, A., Moratti, S., Maestú, F., et al. (2012). Lateral orbitofrontal cortex involvement in initial negative aesthetic impression formation. *PLoS ONE* 7:e38152. doi: 10.1371/journal.pone.0038152

Nadal, M. (2007). *Complexity and Aesthetic Preference for Diverse Visual Stimuli.* PhD thesis. Palma: Universitat de les Illes Balears.

Nadal, M. (2013). The experience of art: Insights from neuroimaging. *Prog. Brain Res.* 204, 135–158. doi: 10.1016/B978-0-444-63287-6.00007-5

Nadal, M., Munar, E., Marty, G., and Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: explaining the divergence of results. *Empir. Stud. Arts* 28, 173–191. doi: 10.2190/EM.28.2.d

Neperud, R. W., and Marschalek, D. G. (1988). Informational and affect bases of aesthetic response. *Leonardo* 21, 305–312. doi: 10.2307/1578660

Nicki, R. M. (1972). Arousal increment and degree of complexity as incentive. *Br. J. Psychol.* 63, 165–171. doi: 10.1111/j.2044-8295.1972.tb02096.x

Nicki, R. M., and Gale, A. (1977). EEG, measures of complexity, and preference for nonrepresentational works of art. *Perception* 6, 281–286. doi: 10.1068/p060281

Normore, L. (1974). "Verbal responses to visual sequences varying in uncertainty level," in *Studies in the New Experimental Aesthetics*, ed. D. E. Berlyne (New York, NY: John Wiley & Sons), 109–119.

North, A. C., and Hargreaves, D. J. (2000). Musical preferences during and after relaxation and exercise. *Am. J. Psychol.* 113, 43–67. doi: 10.2307/1423460

Osborne, J. W., and Farley, F. H. (1970). The relationship between aesthetic preference and visual complexity in absract art. *Psychon. Sci.* 19, 69–70. doi: 10.3758/BF03337424

Palmer, S. E., Schloss, K. B., and Sammartino, J. (2013). Visual aesthetics and human preference. *Annu. Rev. Psychol.* 64, 77–107. doi: 10.1146/annurev-psych-120710-100504

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., et al. (2016). Neuroaesthetics: the cognitive neuroscience of aesthetic experience. *Perspect. Psychol. Sci.* 11, 265–279. doi: 10.1177/1745691615621274

Pelowski, M., Markey, P. S., Lauring, J. O., and Leder, H. (2016). Visualizing the impact of art: an update and comparison of current psychological models of art experience. *Front. Hum. Neurosci.* 10:160. doi: 10.3389/fnhum.2016.00160

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9, 148–158. doi: 10.1038/nrn2317

Petrides, K. V., and Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *J. Appl. Soc. Psychol.* 36, 552–569. doi: 10.1111/j.0021-9029.2006.00019.x

Ramsøy, T. Z., Friis-Olivarius, M., Jacobsen, C., Jensen, S. B., and Skov, M. (2012). Effects of perceptual uncertainty on arousal and preference across different visual domains. *J. Neurosci. Psychol. Econ.* 5, 212–226. doi: 10.1037/a0030198

Reber, R. (2012). "Processing fluency, aesthetic pleasure, and culturally shared taste," in *Aesthetic Science: Connecting Minds, Brains, and Experience*, eds A. P. Shimamura and S. E. Palmer (New York, NY: Oxford University Press), 223–249.

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3

Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Front. Hum. Neurosci.* 9:218. doi: 10.3389/fnhum.2015.00218

Rolls, E. T., and Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Prog. Neurobiol.* 86, 216–244. doi: 10.1016/j.pneurobio.2008.09.001

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cogn. Emot.* 23, 1259–1283. doi: 10.1080/02699930902809375

Russell, P., and George, D. (1990). Relationships between aesthetic response scales applied to paintings. *Empir. Stud. Arts* 8, 15–30. doi: 10.2190/AU1R-6UXE-T14R-04WQ

Sato, K., and Oda, M. (2014). Effect of color and shape on aesthetic evaluation of colored shapes. *Trans. Jpn. Soc. Kansei Eng.* 13, 561–569. doi: 10.5057/jjske.13.561

Schaal, N., Bauer, A.-K., and Müllensiefen, D. (2014). Der Gold-MSI: replikation und validierung eines fragebogeninstrumentes zur messung musikalischer

erfahrenheit anhand einer deutschen stichprobe. *Music. Sci.* 18, 423–447. doi: 10.1177/1029864914541851

Schäfer, T., and Sedlmeier, P. (2011). Does the body move the soul? The impact of arousal on music preference. *Music Percept.* 29, 37–50. doi: 10.1525/mp.2011.29.1.37

Schmitz, G. S., and von Salisch, M. (2002). *Skala zur Erfassung der Emotionalen Selbstwirksamkeit*. Berlin: Freie Universität Berlin.

Schulz, P., Jansen, L. J., and Schlotz, W. (2005). Stressreaktivität: theoretisches konzept und messung [Stress reactivity: theoretical concept and measurement]. *Diagnostica* 51, 124–133. doi: 10.1026/0012-1924.51.3.124

Shimamura, A. P. (2012). "Toward a science of aesthetics," in *Aesthetic Science: Connecting Minds, Brains, and Experience*, eds A. P. Shimamura and S. E. Palmer (New York, NY: Oxford University Press), 3–28.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420

Silvia, P. J. (2005). Emotional responses to art: from collation and arousal to cognition and emotion. *Rev. Gen. Psychol.* 9, 342–357. doi: 10.1037/1089-2680.9.4.342

Stamps Iii, A. E. (2002). Entropy, visual diversity, and preference. *J. Gen. Psychol.* 129, 300–320. doi: 10.1080/00221300209602100

Steyer, R., Schwenkmezger, P., Notz, P., and Eid, M. (1997). *Mehrdimensionaler Befindlichkeitsfragehogen (MDBF)*. Göttingen: Hogrefe.

Storbeck, J., and Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Soc. Personal. Psychol. Compass* 2, 1824–1843. doi: 10.1111/j.1751-9004.2008.00138.x

Tinio, P. P. (2013). From artistic creation to aesthetic reception: the mirror model of art. *Psychol. Aesthet. Creat. Arts* 7, 265–275. doi: 10.1037/a0030872

Vartanian, O., and Goel, V. (2004). Neuroanatomical correlates of aesthetic preference for paintings. *Neuroreport* 15, 893–897. doi: 10.1097/00001756-200404090-00032

Vartanian, O., and Skov, M. (2014). Neural correlates of viewing paintings: evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain Cogn.* 87, 52–56. doi: 10.1016/j.bandc.2014.03.004

Vitz, P. C. (1964). Preferences for rates of information presented by sequences of tones. *J. Exp. Psychol.* 68, 176–183. doi: 10.1037/h0043402

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., et al. (2006). Development of short forms of the empathy quotient (eq-short) and the systemizing quotient (sq-short). *Pers. Individ. Dif.* 41, 929–940. doi: 10.1016/j.paid.2006.03.017

Whitfield, T. W. A. (1983). Predicting preference for familiar, everyday objects: an experimental confrontation between two theories of aesthetic behaviour. *J. Environ. Psychol.* 3, 221–237. doi: 10.1016/S0272-4944(83)80002-4

Wundt, W. (1874). *Grundzüge der Physiologischen Psychologie*. Leipzig: Engelmann.

Yeh, Y. C., Lin, C. W., Hsu, W. C., Kuo, W. J., and Chan, Y. C. (2015). Associated and dissociated neural substrates of aesthetic judgment and aesthetic emotion during the appreciation of everyday designed products. *Neuropsychologia* 73, 151–160. doi: 10.1016/j.neuropsychologia.2015.05.010

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9, 1–27. doi: 10.1080/02699931.2010.497409

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Marin, Lampatz, Wandl and Leder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

2.4.4 Summary and outlook

In Experiment 1, 230 female participants rated either environmental scenes, environmental scenes turned into cartoons or representational paintings presented for 5 s on familiarity, arousal, complexity and one measure of hedonic tone (pleasantness, beauty or liking), which yielded nine experimental conditions. We controlled for participant characteristics such as age, mood prior to the experiment, art interest, trait emotional intelligence, trait empathy and stress reactivity (see Marin & Leder, 2018), which did not differ across the participant groups. The major finding of this experiment was that the multidimensionality of hedonic tone was only present when examining the complexity-hedonic tone relationship in real artworks (i.e., representational paintings). The relationships between one of the three measures of hedonic tone and complexity were similar (and basically not present) for environmental scenes of motivational relevance. This also held true for environmental scenes transformed into cartoons. However, controlling for familiarity, we observed for representational paintings a significant positive relationship between beauty and complexity (in line with Nadal et al., 2010), a non-significant negative relationship between pleasantness and complexity, and no relationship between liking and complexity. Interestingly, the relationship between arousal and complexity was also the strongest for real artworks (nearly $r_s \sim .50$), which replicates findings of Marin and Leder (2013, 2016). Altogether, our data suggests that concepts such as beauty, pleasantness and liking are less correlated during art perception (see also Tiihonen et al., 2022) than during everyday scene perception, which could be a characteristic feature of art perception. In a related study, Sidhu et al. (2018) obtained beauty, liking, complexity, meaningfulness, emotionality and colour warmth ratings in response to representational and abstract paintings. Familiarity was not assessed. They reported that for abstract paintings liking and beauty were much less correlated than for representational paintings, and that in representational paintings complexity correlated much stronger with beauty and liking, respectively, than for abstract art. Subjective complexity predicted only beauty and liking for representational paintings but not for abstract art. Future studies should examine the role of semantic contents, which may partly explain some of the observed differences between abstract and presentational art. Che et al. (2024) reported further evidence for a qualitative difference of liking and beauty when viewing faces and artworks. One can thus conclude that the integration of beauty, pleasantness and liking as different concepts into neuroscientific models of aesthetic experi-

ences (Brattico et al., 2013; Pelowski, Markey, et al., 2017) is justified. Another conclusion regarding Berlyne's (1971) psychobiological model and the reported discrepancy regarding the inverted U-shape relationship between complexity and hedonic tone is that different measures of hedonic tone cannot be deliberately exchanged with each other. Researchers should thus justify the preference for one measure of hedonic tone over another in their respective studies.

In Experiment 2, 92 female participants listened to the 92 excerpts of Romantic solo piano music of Marin and Leder (2013), which were shortened to 15 s because this presentation duration was regarded as sufficient and helped keep the duration of the experiment reasonable. Instead of art interest, we assessed musical sophistication, and all other background variables as well as the procedure were similar as to Experiment 1. Familiarity interacted in different ways with the other subjective measures of interest, again demonstrating how important it is to consider the subject (not only the object) as an important source determining aesthetic appreciation (Berlyne, 1971). We observed more or less no association with complexity, a strong correlation with the measures of hedonic tone, and mostly a moderately negative association with arousal. Overall, this pattern of results is most similar to the one we noted for representational paintings. Interestingly, familiarity and measures of hedonic tone were very similarly associated with each other in all the conditions studied in both experiments, namely around $r_s \sim .60$, which was also reported in Marin and Leder (2013, 2016). In other words, familiarity plays a particular role when measuring hedonic tone across stimulus sets. Considering that the studies were based on a between-subjects design, these findings suggest that the way familiarity is related to complexity, arousal and measures of hedonic tone follows a coherent pattern for each variable of interest that differs regarding the strength (and not regarding the direction) of the relationships across stimulus types: negative with complexity and arousal, positive with hedonic tone. This should be replicated with other stimulus types in the future and integrated into models of aesthetic experiences.

Regarding the predictions made by Berlyne (1971), a strong positive relationship between complexity and arousal was present, replicating the results of Experiments 3 and 4 by Marin and Leder (2013). As for representational paintings, the multidimensionality of hedonic tone became apparent, but the pattern of results differed considerably: Complexity and beauty clearly followed an inverted U-shaped relationship as predicted by Berlyne (positive for paintings), complexity and pleasantness

were weakly positively correlated (negative for paintings and no relationship for females of Experiment 3 in Marin and Leder, 2013), whereas complexity and liking were weakly negatively correlated. The interrelationships between complexity ratings across the three hedonic tone conditions were very high, suggesting that the multidimensionality of hedonic tone cannot be grounded in large differences of perceived complexity across groups. The three groups did not differ in terms of musical sophistication and their general liking for the musical stimuli heard in the experiment, but the participants of the beauty group reported a higher role of music in their lives and a lower amount of music listening than the participants of the other groups.

Altogether, our findings offer further insights into the nature of aesthetic experiences because the multifaceted nature of hedonic tone was absent when rating environmental scenes and cartoon-like scenes. Why the relationship between our measures of hedonic tone and complexity was different when rating representational art versus music still needs to be determined. In such an endeavour it will be necessary to first get a better understanding of musical complexity, especially since Nadal et al. (2010) have shown that the multidimensionality of complexity affects the relationship between beauty and complexity. Furthermore, future research may want to investigate how cartoon-like environmental scenes (here slightly rated above average on their artistic quality and created by using filters and algorithms in Photoshop), and other artificially modified images presented as artworks can be effectively used when studying the nature of aesthetic experiences. It is difficult to find a balance between the control of stimuli regarding, for instance, complexity and semantic contents, and the generation of an artistic surface structure of the stimuli that gives viewers the impression of human-made artworks. Bellaiche et al. (2023) and Ragot et al. (2020) suggest that AI-generated art labelled as human-made is evaluated more highly than artworks labelled as AI-generated. It could be that the cognitive and affective appreciation of artworks is implicitly influenced by knowing that the existence of the images is rooted in human creativity (see also Demmer et al., 2023; Latikka et al., 2023). It is likely that in viewers the current set of cartoons based on environmental scenes did not elicit this impression to a large extent. In other words, cartoons were not strongly appraised as art. Last, our findings could inspire the field of neuroaesthetics (Iigaya et al., 2020), which could systematically study the response to different sets of visual and musical stimuli by asking for several types of hedonic tone ratings. As Brattico et al. (2013) suggests, a co-activation of specific brain regions that are related to beauty

and liking is possible, which actually points to a complex interplay of neuronal activity. In summary, this study demonstrates that a comprehensive and comparative approach (Marin, 2015), something that was already exemplified by Berlyne's *New Experimental Aesthetics*, is a truly fruitful avenue to pursue in the future.

2.5 Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures (Marin & Leder, 2018)

2.5.1 Introduction

Since Fechner and the dawn of empirical aesthetics, aesthetic appreciation has been understood to depend on the characteristics of the perceived object and on the subject's life history (Berlyne, 1971; Bullot & Reber, 2013; de Cruys & Wagemans, 2011; Jacobsen, 2006; Leder et al., 2004; Leder & Nadal, 2014; Pelowski, Markey, et al., 2017; Redies, 2015; Specker et al., 2021, for a review see Pelowski et al., 2016). Here, we analyzed data collected by Marin et al. (2016) with the aim of exploring individual differences (i.e., top-down influences related to personality and expertise/interest) in affective and cognitive processing besides acknowledging the effects of familiarity. We were motivated by the fact that recent research has demonstrated that the complexity-hedonic tone relationship proposed by Berlyne (1971) is affected by the multidimensionality of visual complexity (Nadal et al., 2010), the multifaceted nature of hedonic tone (Marin et al., 2016), by individual differences in the complexity-preference relationship (e.g., Güçlütürk et al., 2016, Street et al., 2016) as well as by expertise (Orr & Ohlsson, 2005).

Experiment 1 involved visual stimuli (environmental scenes, cartoon-like environmental scenes and representational paintings) and Experiment 2 musical stimuli. We specifically explored individual differences related to personality traits when rating subjective arousal and complexity. Due to the fact that our stimuli largely varied in affective contents, we focused on personality traits that play a particular role in affective processing, such as stress reactivity, trait emotional intelligence and empathy. Our advanced statistical linear mixed-effects models also included art interest and musical sophistication, and we again controlled for familiarity effects to enhance comparability across our studies.

2.5.2 Authorship contribution statement

MMM. Conceptualization, methodology, software, validation, formal analysis, investigation, writing − original draft preparation, visualization HL. Supervision, resources, writing − reviewing and editing.

2.5.3 Re-print of the published research article

Personality and Individual Differences 125 (2018) 80–90

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures

Manuela M. Marin[a,b,*], Helmut Leder[b]

[a] Department of Psychology, University of Innsbruck, Maximilianstraße 2, A-6020 Innsbruck, Austria
[b] Department of Basic Psychological Research and Research Methods, University of Vienna, Liebiggasse 5, A-1010 Wien, Austria

ARTICLE INFO

Keywords:
Trait emotional intelligence
Empathy
Expertise
Stress reactivity
Empirical aesthetics
Complexity
Arousal

ABSTRACT

Empirical studies testing Berlyne's (1971) psychobiological model have led to conflicting results, and individual differences have been proposed as a possible explanation. We explored whether top-down influences related to personality may play a role and studied how affect-related personality traits are related to complexity and arousal responses. Two hundred and six females rated either affective environmental scenes, environmental scenes converted into cartoons, or representational paintings (Study 1). Another group of 77 females rated excerpts of piano music (Study 2). We assessed trait emotional intelligence (EI), stress reactivity (SR), empathy (Study 1), emotional self-efficacy (Study 2) as well as engagement with art and music. Linear mixed-effects modelling revealed that affect-related traits emerged as significant predictors in all visual complexity and arousal models, except for the complexity model of environmental scenes. SR predicted arousal induced by environmental scenes, but not by cartoons and paintings, for which an interaction between trait EI and empathy was found. Musical sophistication predicted musical complexity, and the complexity and arousal models comprised interactions between trait EI, SR and emotional self-efficacy. Affect-related traits should be integrated into arousal-based theories of aesthetic experiences, and their impact on perceived complexity and felt arousal varies across stimulus categories.

1. Introduction

Aesthetic experiences are multifaceted phenomena, and current models acknowledge the determining role of the artwork, the characteristics of the perceiver and the socio-cultural context in shaping such experiences. More specifically, several models of art experience have included factors related to the individual, such as personality, prior mood states, as well as expertise (Pelowski, Markey, Lauring, & Leder, 2016). Nevertheless, studying differences within and between individuals still remains a relatively open field of study. Here, our main aim is to identify the role of affect-related traits and expertise in ratings of arousal and complexity, two frequently studied variables of aesthetic experiences, in a group of healthy female university students. We study individual differences in the aesthetic appreciation of music and three sets of affective visual stimuli, thus following the recent trend of a comparative approach across sensory domains (e.g., Marin, 2015).

The notion that studying individual differences in aesthetic experiences may help evaluate the robustness of theories and models has recently gained growing interest among researchers, especially in the context of Berlyne's (1971) psychobiological model. Berlyne describes an inverted U-shaped relationship between so-called collative variables (i.e., familiarity, complexity, novelty, ambiguity etc.) and measures of hedonic tone (i.e., beauty, liking, attractiveness etc.), which is mediated through arousal induced by a stimulus. In this model, stimuli of moderate complexity should be preferred. Some empirical evidence has been adduced in support of this claim, especially regarding the role of arousal in aesthetic experiences (e.g., Marin & Leder, 2013); however, discrepancies regarding the relationship between complexity and measures of hedonic tone have been reported (Chmiel & Schubert, 2017; Marin & Leder, 2013). This relationship may also depend on the underlying dimension of complexity (Nadal, Munar, Marty, & Cela-Conde, 2010) or on the specific measure of hedonic tone (Marin, Lampatz, Wandl, & Leder, 2016). Moreover, musical expertise has been suggested to dissolve (Orr & Ohlsson, 2005) or shift (Hare, 1974; Vitz, 1966) the inverted-U relationship between musical complexity and liking. More recently, Güçlütürk, Jacobs, and Van Lier (2016) showed that individual differences regarding preferences for visual complexity should be considered in the context of Berlyne's psychobiological model, suggesting that across-participant analysis methods may be inadequate to gain a deep understanding of how complexity and liking

* Corresponding author at: Maximilianstraße 2, A-6020 Innsbruck, Austria.
E-mail addresses: manuela.marin@uibk.ac.at (M.M. Marin), helmut.leder@univie.ac.at (H. Leder).

https://doi.org/10.1016/j.paid.2017.12.027
Received 30 August 2017; Received in revised form 14 December 2017; Accepted 18 December 2017
0191-8869/ © 2017 Elsevier Ltd. All rights reserved.

are related. In a similar vein, Street, Forsythe, Reilly, Taylor, and Helmy (2016) reported individual differences in preference for visual complexity evoked by fractals using mixed-effects modelling. We followed this approach by concentrating on two core parameters of Berlyne's influential theory, namely subjective complexity and arousal, and by studying responses to stimuli with affective contents.

In the present study, the issue under scrutiny is how affective (arousal) and cognitive (complexity) processing of visual and musical stimuli can be predicted by individual differences in empathy, emotional intelligence and stress reactivity. These three traits have previously been related to individual differences in affect processing in healthy individuals; therefore, we were interested in examining these traits together in a common framework with regard to aesthetic experiences. More generally, this comprehensive approach of studying more than one personality trait is in line with studies showing how genetic variation of the oxytocin receptor is associated with both empathy and stress reactivity (Rodrigues, Saslow, Garcia, John, & Keltner, 2009), or studies indicating a negative association between trait emotional intelligence and measures of acute and chronic stress level (Mikolajczak, Roy, Luminet, Fillée, & de Timary, 2007; Singh & Sharma, 2012). Correlations between trait EI, empathy and self-efficacy were also reported in an educational context (Kokkinos & Kipritsi, 2012).

Already Mayer, Dipaolo, and Salovey (1990) investigated the role of empathy in the perception of emotional faces, color swatches and abstract designs, showing that participants' ability to perceive emotional contents in visual stimuli was related to the degree of self-reported empathy. The study of empathy recently also gained some attention in the field of visual neuroaesthetics (Freedberg & Gallese, 2007; Schott, 2015), but not much work has been done regarding visual aesthetic experiences (but see Gerger, Pelowski, & Leder, 2017). Empathy has been more frequently studied regarding music-induced emotions. Subcomponents of empathy have been associated with the experience of and preference for sadness in music (e.g., Eerola, Vuoskoski, & Kautiainen, 2016; Garrido & Schubert, 2011). The ability "to feel with someone" generally seems to play a role in processing musical emotions, probably due to an underlying action-perception mechanism (Wöllner, 2012). Taken together, these studies motivated us to further investigate the role of trait empathy, especially in the context of visual aesthetic experiences.

The investigation of trait emotional intelligence (EI), a broader construct than empathy, and its relation to aesthetic experiences may be a worthwhile and timely continuation because both constructs overlap only moderately in healthy individuals (Gökçen, Petrides, Hudry, Frederickson, & Smillie, 2014). Trait EI is a personality trait (Petrides, 2010), related to the Big Five personality traits, with similar genes driving the development of individual differences (Vernon, Villani, Schermer, & Petrides, 2008). A strong overlap between trait EI and the General Factors of Personality has also been reported (Van der Linden, Tsaousis, & Petrides, 2012). Trait EI concerns people's perceptions of their emotional world (i.e., not their actual abilities and skills) and comprises 15 different facets forming factors such as well-being, self-control, emotionality and sociability. The literature on trait EI in the field of empirical aesthetics is scarce, but trait EI has been shown to be positively related to flow experiences in music performance (Marin & Bhattacharya, 2013) and to the amount of musical training (Petrides, Niven, & Mouskounti, 2006). Regarding the aesthetic experience induced by pictures and music varying in emotional content, it is plausible to surmise that arousal responses may change as a function of trait EI.

Another trait associated with individual differences in processing of affective stimuli is stress reactivity. Stress reactivity refers to the "capacity or tendency to respond to a stressor" (Schlotz, 2013, p. 1891) and can be understood as a physiological and psychological response to a perturbation to homeostasis, precipitating the development of cardiovascular diseases and psychosis. Stress reactivity is assumed to be related to high intrusiveness, low self-efficacy, high arousability of the central nervous system and high negative affectivity (Schulz, Jansen, & Schlotz, 2005). In the musical domain, stress-reactivity has been suggested to underlie gender effects observed in psychophysiological arousal responses to music (Nater, Abbruzzese, Krebs, & Ehlert, 2006), with females being more stress reactive than males and showing enhanced psychophysiological arousal responses to unpleasant music. Individual stress reactivity has also been found to be a mediator between habitual music listening-behaviour and indicators of psychological and physiological functioning (Thoma, Scholz, Ehlert, & Nater, 2012). More recently, Gingras, Marin, Puig-Waldmüller, and Fitch (2015) reported that pupil dilation, a measure of emotional arousal, in response to music did not vary as a function of stress reactivity.

Individual differences in art interest and art expertise have been reported to be an important factor in modelling aesthetic experiences (e.g., Istók et al., 2009; Palmer & Griscom, 2013; Silvia, 2013). For instance, Gartus and Leder (2014) studied modern art and reported a positive relationship between interest in modern art and liking among university students, which was not present for general art interest. In a follow-up project, Gartus, Klemer, and Leder (2015) showed that graffiti interest also affected beauty and interest ratings of graffiti art and eye movements. Silvia (2013) demonstrated that people with higher levels of visual art expertise regarded abstract art as more interesting and less confusing. Studying preference for harmony in the visual and musical domains, Palmer and Griscom (2013) found that higher levels of training in the respective art forms were associated with lower preference for harmony, and further, that preference for harmony was generally unrelated to the Big Five personality traits and sensation seeking. van Paasschen, Bacci, and Melcher (2015) showed that art experts and novices reported similar levels of valence and arousal in response to representational and abstract artworks. The results regarding musical training and emotional processing are inconclusive (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005; Mikutta, Maissen, Altorfer, Strik, & König, 2014; Park et al., 2014). Recent research has gone a step further and studied expertise in relation to personality with the motivation to unravel the characteristics of aesthetic people. Musical sophistication, which refers to different aspects of how people engage with music including musical training (Müllensiefen, Gingras, Stewart, & Musil, 2014), was recently reported to be related to the Big 5 personality traits and openness to aesthetics (Greenberg, Müllensiefen, Lamb, & Rentfrow, 2015). Similarly, personality and aesthetic fluency were studied by Fayn, Maccann, Tiliopoulos, and Silvia (2015), who showed that openness to experience and aesthetic fluency were moderately related.

In order not to run into issues of sample size and power, we restricted our sample to female university students, who are frequently studied in the field of empirical aesthetics. Because previous research has shown that traits such as stress reactivity (Nater et al., 2006) and empathy (Kreutz, Schubert, & Mitchell, 2008) can explain gender differences regarding affective responses to music, we decided to explore these variables in more depth among female university students. Moreover, gender differences regarding emotional picture processing are well documented in the literature (Filkowski, Olsen, Duda, Wanger, & Sabatinelli, 2016), with females showing stronger emotional reactions to negative valence and males showing stronger reactions to positive valence (Stevens & Hamann, 2012).

In summary, our research design is characterized by several current trends, namely by considering more than one personality trait (Mikolajczak et al., 2007; Rodrigues et al., 2009), by studying personality traits in relation to interest and sophistication (Greenberg et al., 2015; Palmer & Griscom, 2013), as well as by studying more than one aesthetic domain (Marin, 2015; Marin & Leder, 2013; Palmer & Griscom, 2013; Silvia, Fayn, Nusbaum, & Beaty, 2015). This approach allows us to obtain a more complete picture about the role of individual differences among female university students in the context of Berlyne's psychobiological model, which clearly is a timely issue.

The analyses presented here, employing linear mixed-effects modelling, are based on data collected by Marin et al. (2016), who compared subjective responses to four affective stimulus categories, namely environmental scenes, environmental scenes converted to cartoons, representational paintings (Experiment 1) and solo piano music (Experiment 2). In Experiment 1, the three stimulus types differed regarding their aesthetic quality. By converting photographs of environmental scenes into cartoon-like pictures, having the same semantic and affective contents but different surface structure, a novel, art-like stimulus category was introduced to complement the comparison between environmental scenes and representational paintings. Measures of individual differences regarding art interest, trait EI, trait empathy and stress reactivity were collected by means of standardized questionnaires. In Experiment 2, musical sophistication, trait EI, emotional self-efficacy (emotional competence) and stress reactivity questionnaires were administered.

For the three visual stimulus categories (Study 1), we hypothesized that individual differences regarding the affective traits would more likely emerge with regard to arousal than complexity ratings. This hypothesis is grounded on the assumption that arousal and complexity processing refer to affective and cognitive processing, respectively, and that trait empathy, EI and stress reactivity were previously related to affective processing. Since there is insufficient research on how trait empathy, trait EI and stress reactivity may influence subjective arousal responses when considered together, we refrained from formulating concrete hypotheses. We also surmised that we would observe more indices of individual differences regarding responses to representational paintings than to environmental scenes (Vessel, Stahl, Maurer, Denker, & Starr, 2014). Moreover, we expected that general art interest may not strongly influence arousal and complexity responses (Gartus et al., 2015; Gartus & Leder, 2014; van Paasschen et al., 2015). For music and the influence of traits (Study 2), we hypothesized that the pattern of results for arousal and complexity ratings would be less distinctive due to the fact that arousal and complexity were found to be highly correlated ($r_s \sim 0.80$) for this musical style (Marin & Leder, 2013). Due to inconsistencies in the previous literature on the role of musical background on emotion and complexity ratings of music, we did not formulate a concrete hypothesis about how complexity and arousal ratings may change as a function of musical sophistication.

2. Study1

2.1. Methods

The current research article is based on the analysis of data collected by Marin et al. (2016), and the reader is advised to consult this open access article for detailed information on the participants, materials and procedure. Here, only a summary of the methods of Experiment 1 are given, whereas information on the current statistical analysis is provided in detail.

2.1.1. Participants

Participants were 206 female university students (M = 22.7 years, SD = 3.4) who took part in Experiment 1 as reported in Marin et al. (2016). Sixty-eight participants provided ratings of environmental scenes, 71 ratings of environmental scenes converted into cartoons, and 67 ratings of representational paintings. For each picture category, there were three conditions of hedonic tone rating (beauty, pleasantness and liking), and participants groups were of similar size in each condition after the removal of outliers regarding mood, art interest and personality traits. Table 1 in Marin et al. (2016) offers information on participants' characteristics regarding mean age, trait EI, stress reactivity, empathy quotient and art interest. In general, participants did not differ in these measures across the three levels of hedonic tone per stimulus category. One exception concerns the category of environmental scenes and the average degree of trait EI across the three hedonic tone conditions, where significant differences emerged.

Table 1
Results of linear mixed-effects model predicting complexity ratings of environmental scenes (IAPS pictures).

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	3.674 (0.100)	[3.478, 3.870]			
Random effects					
Familiarity | participants	–	–	2	218.0	< 0.001***
Pictures	–	–	1	1398.0	< 0.001***

Note. Number of observations = 6144, number of stimuli = 96, number of participants = 64; B = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
*** $p < 0.001$.

2.1.2. Materials and procedure

Three different picture sets were studied: affective environmental scenes of different semantic contents (pictures selected from the International Affective Picture System (IAPS), Lang, Bradley, & Cuthbert, 2005), affective environmental scenes (IAPS pictures) converted into cartoons by means of Adobe Photoshop CS5, and representational paintings (see Fig. 1 in Marin et al., 2016). IAPS pictures depicted humans in different social settings as well as animals, food and landscapes. The representational paintings, mostly stemming from the 18th to the beginning of the 20th century, were of similar semantic content. Each set contained 96 pictures, which varied in their emotional content (in arousal and pleasantness, Russell, 1980), as well as in their visual complexity (figure ground composition vs. complex scene), i.e. mostly along the dimension of elements (Marin & Leder, 2016).

In a between-subjects design, participants viewed each picture for 5 s (all pictures were part of the same picture category and presented according to their emotional contents in four randomized blocks). Participants provided ratings of familiarity, felt arousal, complexity and one out of three possible measures of hedonic tone (either beauty, pleasantness or liking). Ratings were given on 7-point Likert scales and self-paced. After the experimental block, participants filled in several standardized questionnaires, which are central to the current analysis regarding individual differences. These questionnaires assessed global trait emotional intelligence (TEIQue-SF, Petrides & Furnham, 2006), stress reactivity (SRS, Schulz et al., 2005), and affective and cognitive empathy (EQ-Short, Wakabayashi et al., 2006). In addition, a self-developed questionnaire to assess art interest was administered (see online Supplementary Table in Marin et al., 2016). For all of these scales only the global score was considered in the statistical analysis.

2.1.3. Statistical analysis

To explore individual differences in subjective complexity and arousal responses, linear mixed-effects models were calculated using R (Version 3.3.0, https://www.r-project.org/) and the *lmer* function in the package lme4 (Bates, Mächler, Bolker, & Walker, 2014) as well as the *anova* function in package car to obtain significance tests (Fox & Weisberg, 2010). Since each picture was rated by each participant, pictures and participants were treated as fully crossed random effects (Baayen, Davidson, & Bates, 2008). Linear mixed-effects models are computed without aggregating over pictures and participants, and more importantly, they allow for the modelling of random intercepts and slopes, which are markers of individual differences.

In each of the three picture sets, the three groups differing in ratings of hedonic tone were aggregated to allow for a large enough sample size. Three models predicting subjective complexity as well as three models predicting felt arousal were built, one for each picture set. Participant, picture and the type of hedonic rating (3 levels) were coded as categorical factors, whereas all other predictors were treated as

covariates and grand mean centered (Enders & Tofighi, 2007). The predictors included familiarity ratings, emotional intelligence, empathy, stress reactivity, art interest (except for IAPS pictures condition, for which art interest was not assessed) as well as type of hedonic tone rating (which was not a significant predictor in any of the models, justifying the aggregation of data across the three levels of hedonic tone).

The random effects structure was carefully determined for each individual model by considering familiarity with the stimuli rated in the experiment. First, we compared a random intercept model, consisting of by-subject random intercepts and by-stimulus random intercepts, with the null model using a likelihood ratio test. For each of the six models, the results showed that the random intercept models were a significant better fit than the respective null models, thus justifying a mixed models approach. Next, the random effects structure was determined by consecutively adding by-subject random slopes for familiarity, as well as by-stimulus random slopes for familiarity. The fit of the random models was again assessed by the likelihood ratio test. The random model with the best fit, leading to convergence of the full model, was chosen for the main analysis. In most cases, models included by-subject random slopes for familiarity and by-stimulus random slopes for familiarity.

To detect influential cases, Cook's distances (Cook & Weisberg, 1982) were calculated for a full model that included all main effects and 2- and 3-way interactions of interest. Influential cases were identified as having a value larger than $4/N\text{-}k\text{-}1$, with k corresponding to the number of predictors. After influential cases were removed, the model was simplified by using the *step* function of the lmerTest package, which performs backward elimination of non-significant effects of linear mixed models. Lastly, all models were compared to the models containing only random effects by the likelihood ratio test. Regression diagnostics were performed by visually inspecting the plots of residuals for normality and homogeneity. *P*-values were calculated using likelihood ratio tests (Barr, Levy, Scheepers, & Tily, 2013) and Type 2 sums of squares. Interaction diagrams (see Supplementary File) were made in R using the interplot package (Solt & Hu, 2015, https://cran.r-project.org/web/packages/interplot/interplot.pdf) and the effects package (Fox et al., 2016, ftp://cran.r-project.org/pub/R/web/packages/effects/effects.pdf).

2.2. Results

Cronbach alphas were calculated to assess the reliability of the scales, and they were considered as acceptable in all cases (ranging between $\alpha = 0.76$–0.87). The intercorrelations between these measures are presented in Supplementary Table A.1 and given separately for each participant group. In all three groups, the correlation between trait EI and empathy was positive with a strong effect size ($r \sim 0.50$). Furthermore, the relationship between trait EI and stress reactivity was negative and of a slightly lower effect size ($r \sim -0.40$). We did not observe any other significant correlations or indications of multicollinearity between the variables. Cartoons ($M = 3.81$, $SD = 0.58$) were rated as more complex than IAPS pictures ($M = 3.69$, $SD = 0.73$), $t(95) = -3.49$, $p = 0.001$, and cartoons ($M = 4.0$, $SD = 0.72$) were rated as less arousing than IAPS pictures ($M = 4.25$, $SD = 0.80$), $t(95) = 5.13$, $p < 0.001$.

2.2.1. Individual differences in subjective complexity responses

A linear mixed-effects model was built to predict subjective complexity ratings of environmental scenes ($M = 3.69$, $SD = 0.73$, based on $N = 68$ participants), using the following predictors as fixed effects: trait EI, empathy, stress reactivity, familiarity as well as a variable specifying the group of hedonic tone (3 levels). Interaction terms comprised three 2-way interactions between trait EI, empathy and stress reactivity as well as the 3-way interaction between these variables. The best fitting random effects structure included by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts. After excluding four influential cases, the results of a backward elimination procedure (Table 1) revealed no significant fixed effects, but the random effects were highly significant.

Table 2
Results of linear mixed-effects model predicting complexity ratings of environmental scenes converted into cartoons.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	3.805 (0.090)	[3.629, 3.982]			
Stress reactivity	0.028 (0.008)	[0.012, 0.043]	1	11.77	< 0.001***
Random effects					
Familiarity | participants	–	–	2	470.7	< 0.001***
Familiarity | pictures	–	–	2	13.9	< 0.001***

Note. Number of observations = 6720, number of stimuli = 96, number of participants = 70; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
*** $p < 0.001$.

The linear mixed-effects model predicting complexity ratings of environmental scenes (IAPS pictures) converted into cartoons ($M = 3.81$, $SD = 0.58$, based on $N = 71$) was built in a similar way, but in addition included art interest as a fixed effect. Therefore, the interaction terms comprised six 2-way interactions, as well as one 3-way interaction (between stress reactivity, trait EI and empathy) and one 4-way interaction (between stress reactivity, trait EI, empathy and art interest). The best-fitting random effects structure included by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts and slopes for familiarity. After excluding one influential case, the results (Table 2) indicated a significant main effect of stress reactivity ($b = 0.028$, $SE = 0.008$, $\chi^2(1) = 11.77$, $p < 0.001$), suggesting that increased stress reactivity is associated with a higher degree of perceived complexity. No other significant main effects or interactions were observed. The random effects were highly significant, and the final model was a better fit than the random effects model, $\chi^2(1) = 11.77$, $p < 0.001$.

The third model was built to predict complexity ratings of representational paintings ($M = 4.10$, $SD = 0.69$, based on $N = 67$) and comprised the same factors and interactions as the model computed for complexity ratings of cartoons. The best-fitting random effects structure included by-subject random intercepts and slopes for familiarity and by-

Table 3
Results of linear mixed-effects model predicting complexity ratings of representational paintings.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	4.073 (0.104)	[3.870, 4.275]			
Familiarity	0.081 (0.021)	[0.040, 0.122]	1	14.61	< 0.001***
Stress reactivity	0.009 (0.010)	[−0.009, 0.028]	1	0.99	0.319
Empathy	0.008 (0.002)	[−0.016, 0.031]	1	0.41	0.522
Stress reactivity × empathy	0.004 (0.002)	[0.001, 0.007]	1	6.11	0.013*
Random effects					
Familiarity | participants	–	–	2	57.3	< 0.001***
Familiarity | pictures	–	–	2	12.6	0.002**

Note. Number of observations = 6144, number of stimuli = 96, number of participants = 64; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

stimulus random intercepts and slopes for familiarity. Three influential cases were detected and removed. The results (Table 3) showed a significant effect of familiarity ($b = 0.081$, $SE = 0.021$, $\chi^2(1) = 14.61$, $p < 0.001$), suggesting that more familiar pictures were associated with a higher degree of perceived complexity. There was also a significant interaction between stress reactivity and empathy ($b = 0.004$, $SE = 0.002$, $\chi^2(1) = 6.11$, $p = 0.013$). The random effects were highly significant, and the final model was a better fit than the random effects model, $\chi^2(4) = 23.65$, $p < 0.001$. Supplementary Fig. B.1A shows changes in the coefficient of stress reactivity in the two-way interaction conditional on the degree of empathy. The plot illustrates that with increasing empathy, the magnitude of the coefficient of stress reactivity on complexity ratings also increases. Supplementary Fig. B.1B shows how empathy and stress reactivity determine complexity ratings: among participants with low empathy, an increase in stress reactivity is associated with lower complexity ratings, whereas among those with high empathy, higher levels of stress reactivity are associated with an increase in complexity ratings.

2.2.2. Individual differences in subjective arousal responses

For the linear mixed-effects model predicting arousal ratings of environmental scenes ($M = 4.25$, $SD = 0.80$, based on $N = 68$) (Table 4), the best-fitting random structure consisted of by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts and slopes for familiarity. Trait EI, empathy, stress reactivity, familiarity ratings as well as a variable specifying the group of hedonic tone (3 levels) were included as fixed effects. Interaction terms comprised three 2-way interactions between trait EI, empathy and stress reactivity as well as the 3-way interaction between these variables. After the computation of the full model, Cook's distances were calculated and three influential cases were detected and removed. The final model was fitted using a backward elimination procedure and revealed main effects of familiarity ($b = 0.185$, $SE = 0.027$, $\chi^2(1) = 37.64$, $p < 0.001$) and stress reactivity ($b = 0.029$, $SE = 0.009$, $\chi^2(1) = 9.65$, $p = 0.002$). Two higher-order significant interactions were observed, namely between stress reactivity and empathy ($b = 0.003$, $SE = 0.002$, $\chi^2(1) = 4.94$, $p = 0.026$), and between trait EI and empathy ($b = 0.002$, $SE = 0.001$, $\chi^2(1) = 6.64$, $p = 0.010$). All random effects were highly significant, $p < 0.001$. The final model also yielded a better fit than the random effects model, $\chi^2(6) = 58.85$, $p < 0.001$.

Table 4
Results of linear mixed-effects model predicting arousal ratings of environmental scenes (IAPS pictures).

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	4.095 (0.117)	[3.869, 4.322]			
Familiarity	0.185 (0.027)	[0.131, 0.239]	1	37.64	< 0.001***
Stress reactivity	0.029 (0.009)	[0.011, 0.047]	1	9.65	0.002**
Trait EI	−0.005 (0.006)	[−0.018, 0.008]	1	0.58	0.448
Empathy	−0.002 (0.015)	[−0.030, 0.026]	1	0.01	0.905
Stress reactivity × empathy	0.003 (0.002)	[0.0004, 0.006]	1	4.94	0.026*
Trait EI × empathy	0.002 (0.001)	[0.001, 0.004]	1	6.64	0.010*
Random effects					
Familiarity | participants	–	–	2	215.1	< 0.001***
Familiarity | pictures	–	–	2	30.3	< 0.001***

Note. Number of observations = 6240, number of stimuli = 96, number of participants = 65; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 5
Results of linear mixed-effects model predicting arousal ratings of environmental scenes converted into cartoons.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	4.175 (0.112)	[3.957, 4.394]			
Familiarity	0.092 (0.029)	[0.035, 0.148]	1	9.79	0.002**
Trait EI	−0.005 (0.006)	[−0.017, 0.006]	1	0.83	0.362
Empathy	−0.008 (0.015)	[−0.037, 0.022]	1	0.27	0.606
Trait EI × empathy	−0.002 (0.001)	[−0.004, −0.001]	1	6.51	0.011*
Random effects					
Familiarity | participants	–	–	2	316.0	< 0.001***
Familiarity | pictures	–	–	2	20.2	< 0.001***

Note. Number of observations = 6432, number of stimuli = 96, number of participants = 67; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 6
Results of linear mixed-effects model predicting arousal ratings of representational paintings.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	3.894 (0.107)	[3.686, 4.102]			
Familiarity	0.093 (0.030)	[0.035, 0.152]	1	9.59	0.002**
Trait EI	0.001 (0.005)	[−0.009, 0.011]	1	0.03	0.854
Empathy	0.007 (0.014)	[−0.021, 0.035]	1	0.27	0.607
Trait EI × empathy	−0.001 (0.001)	[−0.003, −0.0002]	1	4.93	0.026*
Random effects					
Familiarity | participants	–	–	2	156.0	< 0.001***
Pictures	–	–	1	967.0	< 0.001***

Note. Number of observations = 5856, number of stimuli = 96, number of participants = 61; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Supplementary Fig. B.2A–D contains a visualization of the two interaction effects. The findings regarding the interaction between stress reactivity and empathy (Fig. B.2B) indicate that, among female university students with low stress reactivity, an increase in empathy is associated with a decrease in arousal ratings. Among females with high stress reactivity, an increase in empathy is associated with an increase in arousal. The interaction between trait EI and empathy (Fig. B.2D) was of a similar nature: Among females with low empathy, those with low trait EI reported increased arousal, whereas among females with high empathy, those with higher levels of trait EI reported increased arousal in response to environmental scenes.

The best-fitting random effects structure of the model predicting arousal ratings of environmental scenes converted into cartoons ($M = 4.0$, $SD = 0.72$, based on $N = 71$) (Table 5) consisted of by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts and slopes for familiarity. The full model included the following predictors: trait EI, empathy, stress reactivity, art interest, hedonic tone and familiarity. Furthermore, six 2-way interactions

between trait EI, empathy, stress reactivity and art interest as well as one 3-way interaction (between trait EI, empathy and stress reactivity) and one 4-way interaction (between trait EI, empathy, stress reactivity and art interest) were added to the model. Four influential cases were detected and removed from the data set. Following a backward elimination procedure, the model with the best fit included significant main effects of familiarity ($b = 0.092$, $SE = 0.029$, $\chi^2(1) = 9.79$, $p = 0.002$), as well as a significant interaction (Fig. S4) between trait EI and empathy ($b = -0.002$, $SE = 0.001$, $\chi^2(1) = 6.51$, $p = 0.011$). All random effects were highly significant, $p < 0.001$. We did not observe any other significant main effects and interactions. The final model provided a better fit than the random effects model, $\chi^2(4) = 17.67$, $p = 0.001$. Fig. B.3A–B shows the negative interaction predicting arousal ratings of cartoons. Among female university students with low empathy, an increase in trait EI was associated with higher arousal responses. Among females with high empathy, an increase in trait EI was associated with a decrease in arousal.

Finally, arousal ratings of representational paintings were predicted ($M = 3.81$, $SD = 0.66$, based on $N = 67$) (Table 6). The best-fitting random effects structure consisted of by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts. The predictors of the full model included main effects of trait EI, empathy, stress reactivity, art interest, hedonic tone as well as familiarity. Six 2-way interactions between trait EI, empathy and stress reactivity were included as well as one 3-way interaction (between trait EI, empathy and stress reactivity) and one 4-way interaction (between trait EI, empathy, stress reactivity and art interest). Six influential cases were detected and removed. The backward elimination procedure yielded a final model including one main effect of familiarity ($b = 0.093$, $SE = 0.030$, $\chi^2(1) = 9.59$, $p = 0.002$) and a significant interaction between trait EI and empathy ($b = -0.001$, $SE = 0.001$, $\chi^2(1) = 4.93$, $p = 0.026$). No other significant main effects or interactions were observed. All random effects were highly significant, $p < 0.001$. The final model also provided a better fit than the model containing only random effects, $\chi^2(4) = 15.88$, $p = 0.003$. The interaction diagrams (Fig. B.4A–B) for the negative interaction between trait EI and empathy revealed that among female university students with low trait EI, an increase in empathy was associated with an increase in arousal. The difference between females with low and high empathy was less pronounced among females with high trait EI.

2.3. Discussion

We explored individual differences regarding the experience of affective visual stimuli varying in their aesthetic quality, comprising environmental scenes, environmental scenes converted into cartoons, and representational paintings. To be specific, we investigated how trait EI, empathy, stress reactivity and general art interest may be related to complexity and arousal ratings of female university students, respectively. Linear mixed-effects models accounting for self-reported familiarity revealed that random effects due to individual differences were significantly present in all models, indicating that mixed-effects modelling is justified in the current context. In general, these findings are in line with previous work by Güçlütürk et al. (2016) and Street et al. (2016) who demonstrated individual differences with regard to visual complexity judgments.

Our hypothesis that complexity ratings were less related to affective traits than arousal ratings was partly supported by the data. The model predicting complexity ratings of environmental scenes did not include any significant fixed effects. However, the complexity model computed for environmental scenes converted into cartoons included stress reactivity as a significant positive predictor. Stress reactivity also interacted with trait empathy when modelling complexity ratings of representational paintings, for which also familiarity emerged as a significant positive predictor. Because the current stimulus sets varied in complexity as well as in semantic and affective contents, our results may suggest that cognitive processes related to complexity judgments may be influenced by individual differences normally found in affective processing. This speculation is spurred by findings showing that affective information can impact on cognitive processes (Blanchette & Richards, 2010; Clore & Huntsinger, 2007). Moreover, neurophysiological studies on affective processing of IAPS pictures have not provided a clear picture on how affective processing interacts with the processing of physical characteristics, such as complexity: some studies have lend support to the idea that complexity is related to early but not late event-related potentials during affective picture processing (Bradley, Hamby, Löw, & Lang, 2007). More recently, it has been suggested that physical characteristics influence early and late processing stages in different ways (Wiens, Sand, & Olofsson, 2011).

As predicted, we also observed individual differences regarding arousal ratings, and significant random effects emerged for all three models. Stress reactivity only appeared as a significant predictor in the model of environmental scenes, suggesting that stress reactivity may come into play when arousal is induced by scenes of motivational relevance (e.g., Bradley, Codispoti, Sabatinelli, & Lang, 2001) and not by stimuli of a more artistic quality. However, previous research has shown that stress reactivity also predicted psychophysiological arousal responses to unpleasant music in females (Nater et al., 2006). Therefore, we need to be careful and refrain from jumping to the tempting conclusion that stress reactivity may only come into play when female university students respond to stimuli of clear evolutionary significance. Moreover, we should note that in the current study stress reactivity also predicted complexity ratings of environmental scenes turned into cartoons and paintings.

Modelling arousal ratings in response to cartoons revealed a significant positive interaction between trait EI and empathy, which was, importantly, also present in the model computed for representational paintings. Conversely, trait EI and empathy interacted negatively with each other in the arousal model of environmental scenes. This complex interplay between two moderately related concepts demonstrates the subtlety of individual differences in arousal responses, warranting further investigation of the role of empathy and emotion related individual differences during the perception of visual aesthetic stimuli (Freedberg & Gallese, 2007; Schott, 2015). One pathway to elucidate the underlying nature of the trait EI – empathy interaction may be to study how the subdimensions of global trait EI determine this interaction effect. For this purpose, the full form of the TEIQue should be used rather than the short form of this measure due to its higher internal consistency. For instance, this line of research could build upon a study showing how subdimensions of trait EI predict affective reactions to films in female university students (Papousek, Freudenthaler, & Schulter, 2008). We suggest to tackle this endeavour by also studying abstract artworks which do not depict pain, faces and emotional scenes.

In line with our hypothesis, general art interest was not a significant predictor of complexity ratings among university students. This finding is in line with Gartus et al. (Gartus et al., 2015, Gartus & Leder, 2014), who showed that only specific art interest in relation to the art style under investigation may lead to differences in measures of aesthetic experience. Silvia (2013) found that higher degrees of art expertise (i.e., aesthetic sensitivity) were associated with less confusion regarding abstract art. However, the discrepant results may be explained by the fact that we investigated responses to three representational picture sets that contained semantic and affective meaning, and did not assess art expertise with the same measure as Silvia did. Moreover, we studied a participant sample that is commonly studied in the field, but which usually does not comprise experts in visual art or artists. Finally, general art interest did not predict arousal ratings, which is in line with findings reported by van Paasschen et al. (2015), suggesting that emotional responses may be rather unaffected by interest and expertise.

3. Study 2

3.1. Methods

As in Study 1, the current data analysis is based on data collected by Marin et al. (2016) in Experiment 2, and the reader is advised to consult this article for detailed information on the participants, materials and procedure. The procedure for the statistical analysis (i.e., linear mixed-effects modelling using R) was identical to the one reported in Study 1. We included emotional self-efficacy (Saarni, 1999) because it partly overlaps with trait EI on dimensions such as emotion perception, expression and empathy. At the same time, this measure of emotional self-efficacy constitutes a somewhat narrower concept than trait EI, the latter additionally comprising trait happiness, optimism, impulsiveness and assertiveness. Another difference between these measures is that trait EI is treated within the context of personality, whereas emotional self-efficacy is regarded as a measure of emotion regulation and competence, which has been linked to self-esteem and moral dispositions (Saarni, 1999). However, both concepts, trait EI and emotional self-efficacy, have been linked to resilience, well-being and mental health (Martins, Ramalho, & Morin, 2010).

3.1.1. Participants

Participants were 77 female university students (all non-musicians, $M = 21.1$ years, $SD = 2.0$) who took part in Experiment 2 as described in Marin et al. (2016). These participants rated excerpts of Romantic piano music in three groups, varying in the type of hedonic tone rating (beauty, pleasantness and liking). These three groups did not significantly differ regarding their mood prior to the experiment, musical sophistication, trait EI, stress reactivity as well as emotional self-efficacy (see Table 6 in Marin et al., 2016).

3.1.2. Materials and procedure

The musical stimuli consisted of 92 excerpts of Romantic piano solo music, taken from Marin and Leder (2013) and cut to a length of 15 s. They were selected from a wide range of works by representative composers of that period, such as Chopin, Brahms, Mendelssohn, Grieg and Schumann. These excerpts varied in their emotional contents (ranging from low to high arousal and from low to high pleasantness), covering both dimensions of the arousal-pleasantness emotion space (Russell, 1980). Excerpts were presented in four blocks according to their emotional contents and rated on 7-point rating scale for familiarity, complexity, felt arousal and one measure of hedonic tone (either beauty, pleasantness or liking). The questionnaires administered after the experimental block comprised, as in Experiment 1, the TEIQue-SF and SRS, but also the emotional self-efficacy (ES) scale by Schmitz and von Salisch (2002) and the self-report measure of the Goldsmiths Musical Sophistication Index (Gold-MSI, Schaal, Bauer, & Muellensiefen, 2014).

3.2. Results

Cronbach's alphas for the scales were computed, revealing values to be considered as acceptable: $\alpha_{\text{trait EI}} = 0.83$, $\alpha_{ES} = 0.86$, and $\alpha_{\text{stress}} = 0.85$. Next, Pearson product-moment correlations between these scales were computed (Supplementary Table A.2). The correlation between trait EI and emotional self-efficacy was strong ($r = 0.69$). As expected, the correlation between trait EI ($r = -0.37$) and stress reactivity as well as the one between emotional self-efficacy and stress reactivity ($r = -0.39$) were negative. The positive association between trait EI and musical sophistication was not significant ($r = 0.16$). No other notable correlations were present.

A linear mixed-effects model predicting complexity ratings of musical excerpts ($M = 4.37$, $SD = 0.90$), based on $N = 77$ participants) was fitted to the data (Table 7). The best random effects structure involved by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts. The full model comprised familiarity, stress reactivity, trait EI, emotional self-efficacy (ES), and musical sophistication as fixed factors. Moreover, six two-way interactions between these factors were added, as well as one 3-way interaction (between stress reactivity, trait EI, and emotional self-efficacy) and one 4-way interaction (between stress reactivity, trait EI, emotional self-efficacy and musical sophistication). After the removal of six influential cases, a backward elimination procedure revealed a significant effect of familiarity ($b = 0.060$, $SE = 0.019$, $\chi^2(1) = 9.37$, $p = 0.002$) and musical sophistication ($b = 0.021$, $SE = 0.005$, $\chi^2(1) = 14.61$, $p < 0.001$), as well as a significant 3-way interaction between stress reactivity, trait EI and emotional self-efficacy ($b = -0.0001$, $SE = 0.00004$, $\chi^2(1) = 4.99$, $p = 0.026$). The interactions between stress reactivity and trait EI ($b = -0.002$, $SE = 0.001$, $\chi^2(1) = 8.09$, $p = 0.004$) and between stress reactivity and emotional self-efficacy ($b = 0.002$, $SE = 0.001$, $\chi^2(1) = 4.47$, $p = 0.035$) were also significant. This final model performed better than the model including only random effects, $\chi^2(9) = 36.91$ $p < 0.001$. Supplementary Fig. B.5A–D illustrates the nature of the interaction between stress reactivity and trait EI and stress reactivity and emotional self-efficacy, respectively. The negative interaction between stress reactivity and trait EI (Fig. B.5B) implies that among female university students with low stress reactivity, complexity responses increase with level of trait EI, whereas among females who score high on stress reactivity, complexity responses decrease with increasing level of trait EI. The positive interaction between stress reactivity and emotional self-efficacy (Fig. B.5D) implies that, among females with high emotional self-efficacy, complexity ratings increased with level of stress reactivity, whereas stress reactivity did not strongly impact on complexity ratings among females with low to moderate emotional self-efficacy.

The mixed-effects model predicting arousal ratings of piano solo music ($M = 3.87$, $SD = 0.83$, based on $N = 77$) was computed in a manner similar to the one for complexity ratings, with a full model involving the same factors and interactions (Table 8). Again, the best fitting random structure involved by-subject random intercepts and slopes for familiarity and by-stimulus random intercepts. After the removal of seven influential cases, the final model revealed a significant effect of familiarity ($b = 0.087$, $SE = 0.027$, $\chi^2(1) = 9.84$, $p = 0.002$), an interaction between stress reactivity and trait EI ($b = -0.003$, $SE = 0.001$, $\chi^2(1) = 11.63$, $p < 0.001$), an interaction between stress reactivity and emotional self-efficacy ($b = 0.003$, $SE = 0.001$, $\chi^2(1) = 5.19$, $p = 0.023$) as well as an interaction between trait EI and emotional self-efficacy ($b = -0.001$, $SE = 0.0004$, $\chi^2(1) = 5.52$, $p = 0.019$). The final model performed better than the model including only random effects, $\chi^2(7) = 27.51$ $p < 0.001$. Supplementary Fig. B.6A–F contains the interaction diagrams of the three significant interactions. Fig. B.6B illustrates the negative interaction between stress reactivity and trait EI. Among female university students with low trait EI, arousal increased with increasing level of stress reactivity, whereas among females with high trait EI arousal decreased with increasing stress reactivity. In Fig. B.6D, the interaction between stress reactivity and emotional self-efficacy is shown. Among females with low emotional self-efficacy, reported arousal decreased with level of stress reactivity, while among those with high emotional self-efficacy arousal increased. Finally, Fig. B.6F represents the negative interaction between trait EI and emotional self-efficacy. Among females with low emotional self-efficacy, arousal increased with level of trait EI, whereas among females with high emotional self-efficacy arousal decreased with increasing level of trait EI.

3.3. Discussion

We examined how musical complexity and arousal are related to affect-related traits and musical sophistication by controlling for familiarity effects. Our results yielded a significant effect of familiarity and indices of individual differences as shown by significant by-subject

Table 7
Results of linear mixed-effects model predicting complexity ratings of piano solo music.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	4.347 (0.119)	[4.119, 4.575]			
Familiarity	0.060 (0.019)	[0.022, 0.096]	1	9.37	0.002**
Stress reactivity	0.011 (0.010)	[−0.008, 0.029]	1	1.31	0.252
Trait EI	−0.001 (0.005)	[−0.011, 0.010]	1	0.02	0.886
ES	−0.016 (0.009)	[−0.033, 0.002]	1	3.04	0.081
Gold MSI	0.021 (0.005)	[0.011, 0.031]	1	14.61	< 0.001***
Stress reactivity × Trait EI	−0.002 (0.001)	[−0.003, −0.001]	1	8.09	0.004**
Stress reactivity × ES	0.002 (0.001)	[0.0002, 0.004]	1	4.47	0.035*
Trait EI × ES	−0.00004 (0.0004)	[−0.0009, 0.0008]	1	0.01	0.923
Stress reactivity × Trait EI × ES	−0.0001 (0.00004)	[−0.0002, −0.00001]	1	4.99	0.026*
Random effects					
Familiarity | participants	–	–	2	84.5	< 0.001***
Musical excerpts	–	–	1	3235.8	< 0.001***

Note. Number of observations = 6532, number of stimuli = 92, number of participants = 71; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test; ES = Emotional self-efficacy.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

random intercepts and slopes for familiarity. Furthermore, the data provide convincing evidence for personality-related individual differences among female university students by demonstrating that both complexity and arousal ratings of piano solo music excerpts changed as a function of stress reactivity, trait EI and emotional self-efficacy. In fact, the pattern of results observed for complexity and arousal ratings is very similar, thus generally in line with our hypothesis, but there were also some notable differences. For example, musical sophistication emerged as a significant predictor for complexity, whereas arousal ratings were unaffected by differences in musical sophistication. In addition, the complexity model also included a significant three-way interaction between stress reactivity, trait EI and emotional self-efficacy, which was not present in the arousal model. Instead, the arousal model comprised a significant interaction between trait EI and emotional self-efficacy, which was not present in the complexity model. In both models, we did not find significant main effects of stress reactivity, trait EI and emotional self-efficacy, but only significant interactions between these measures.

Our data supports the view that stress reactivity plays a role in emotional responses to music in females, even in the case of a rather pleasant musical style in comparison to heavy metal music studied by Nater et al. (2006). Stress reactivity did not emerge as a main effect if considered together with trait EI and emotional self-efficacy, which suggests that it is worthwhile to differentiate and consider different affect-related traits. Moreover, we observed a significant interaction between trait EI and emotional self-efficacy supporting this view. A closer look at the interaction between stress reactivity and trait EI for complexity and arousal ratings revealed that among female university students with high trait EI, ratings decreased with increasing levels of stress reactivity. This pattern was reversed for females scoring low on trait EI, for who ratings increased with stress reactivity. In contrast, for females with low emotional self-efficacy, an increase of stress reactivity resulted in lower ratings, and among those with high emotional self-efficacy, ratings increased with an increase of stress reactivity. On the one hand, there is a considerable overlap between trait EI and emotional self-efficacy ($r = 0.69$), whereas on the other hand, several differences in sub-dimensions of these scales may explain the different nature of the two interactions with stress reactivity. Our findings suggest a need for greater differentiation between these measures, especially since trait EI is often named emotional self-efficacy as well (Petrides, 2010).

We observed that with increasing musical sophistication music was rated as more complex. Interestingly, this effect was not present for arousal ratings. Given that self-reported musical sophistication is associated with higher perceptual abilities regarding music (Müllensiefen et al., 2014), our findings would support the view that female university students with higher musical sophistication may be better able to perceive the inherent intricacies and details of the musical structure. This interpretation is against the prevailing conception that increased experience reduces perceived complexity (Orr & Ohlsson, 2005).

Table 8
Results of linear mixed-effects model predicting arousal ratings of piano solo music.

Fixed effects	*B (SE)*	95% CI	*df*	Chi^2	*p(LRT)*
(Intercept)	3.977 (0.109)	[3.765, 4.187]			
Familiarity	0.087 (0.027)	[0.033, 0.141]	1	9.84	0.002**
Stress reactivity	−0.008 (0.009)	[−0.026, 0.010]	1	0.82	0.367
Trait EI	0.004 (0.005)	[−0.005, 0.013]	1	0.71	0.400
ES	−0.010 (0.008)	[−0.026, 0.007]	1	1.26	0.263
Stress reactivity × Trait EI	−0.003 (0.001)	[−0.004, −0.001]	1	11.63	< 0.001***
Stress reactivity × ES	0.003 (0.001)	[0.0004, 0.005]	1	5.19	0.023*
Trait EI × ES	−0.001 (0.0004)	[−0.002, −0.0001]	1	5.52	0.019*
Random effects					
Familiarity | participants	–	–	2	230	< 0.001***
Musical excerpts	–	–	1	2434	< 0.001***

Note. Number of observations = 6440, number of stimuli = 92, number of participants = 70; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test; ES = Emotional self-efficacy.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

4. General discussion

Within the framework of Berlyne's (1971) psychobiological model of aesthetic experiences, researchers have become increasingly interested in the study of individual differences (e.g., Güçlütürk et al., 2016; Orr & Ohlsson, 2005). This research project examined whether and how affect-related traits are manifested in complexity and arousal ratings of affective visual and musical stimuli by following the current trend of using linear mixed-effects modelling (Fayn et al., 2015; Gartus et al., 2015; Silvia, 2005; Street et al., 2016). In line with our predictions, we observed that affect-related traits particularly predicted arousal ratings and, to a much lesser degree, complexity judgements of visual stimuli, whereas for music both arousal and complexity ratings were affected by these traits. To be specific, our data provides convincing evidence that arousal induced by affective environmental scenes changed as a function of trait EI, stress reactivity and empathy, whereas environmental scenes converted into cartoons, as well as representational paintings changed as a function of trait EI and empathy. Moreover, we were able to demonstrate that arousal and complexity ratings of piano solo music were also influenced by trait EI, emotional self-efficacy and stress reactivity. These findings clearly indicate that affect-related traits need to be incorporated into Berlyne's (1971) theory, or possibly into any other arousal-based theory of aesthetic experience, especially if affective stimuli are being investigated. More generally, our data also suggest that individual differences need to be systematically investigated in relation to all core variables of Berlyne's theory because the traits under investigation may – or may not – have an impact on specific variables, and importantly, the patterns of results may differ across sensory domains and stimulus types. Future studies will have to concentrate on determining how the complexity-arousal-hedonic tone relationship, as described by Berlyne, is being moderated or mediated by these and other types of traits.

In Berlyne's (1971) model, the motivational condition of curiosity as well as exploratory behaviour constitute important components, whose relationship with collative properties was not investigated in the present studies. Therefore, we suggest to expand our approach by measuring other dependent variables, such as looking or listening time, questions regarding the stimuli and ratings of interest (see, e.g., Berlyne, 1974; Saklofske, 1975a, 1975b). Studying interest is also of relevance for more recent theories which link cognitive appraisals with aesthetic emotions (Silvia, 2005) and which also explore individual differences (e.g., Fayn et al., 2015; Nishikawa & Amemiya, 2017).

Our results also revealed how affect-related traits are correlated among female university students. Zero-order correlations between trait EI and stress reactivity were of a moderate and negative effect size in all groups of participants in both studies, which replicates earlier reports (Mikolajczak et al., 2007; Singh & Sharma, 2012). In Study 1, trait EI and empathy were strongly positively correlated across groups (see Gökçen et al., 2014), whereas stress reactivity and empathy were not correlated at all. In Study 2, trait EI and emotional self-efficacy were strongly correlated, and stress reactivity was also negatively correlated with emotional self-efficacy (Schulz et al., 2005). In both studies, these affect-related traits were not significantly correlated with measures of interest and expertise. Thus, these results may present a first step of describing the affective nature of healthy female university students, which may generalize to other populations.

In our sample of female university students, general art interest was not found to be a predictor of complexity and arousal ratings of visual stimuli, but musical sophistication played a role in musical complexity judgments. One possible interpretation of this finding is based on research by Gartus et al. (Gartus et al., 2015, Gartus & Leder, 2014), who showed that general art interest is not a reliable predictor of eye movements and subjective ratings of graffiti art, while graffiti interest was. We suggest that art interest as well as interest with respect to the style under investigation need to be assessed more comprehensively in the future, also by including participants with very high art interest. On the other hand, it is also possible that affective responses to visual stimuli are not largely driven by factors of expertise and interest (van Paasschen et al., 2015). Furthermore, it is likely that musical sophistication and visual art interest, as currently assessed, are not measuring the same underlying concepts. The former was measured by means of a standardized questionnaire covering a myriad of facets related to musical sophistication (Müllensiefen et al., 2014). The art interest questionnaire, however, has not been developed along the same dimensions (see Marin et al., 2016 for details on the items) and emotional engagement with art was not assessed. Future research will have to meet the challenge of developing measures that are able to assess aspects of artistic or aesthetic sophistication in a way that comprises interest as well as expertise.

Our measures of interest and expertise did not significantly correlate or interact with any affect-related personality traits. This is not in line with research showing that musical sophistication is associated with openness to experience (Müllensiefen et al., 2014), or more specifically with openness to aesthetics (Greenberg et al., 2015). Petrides et al. (2006) also reported a correlation between amount of musical training and trait EI. Furthermore, Fayn et al. (2015) reported that aesthetic fluency, another measure of expertise, was moderately associated with openness to experience as well as to pleasure and interest induced by modern artworks. It thus remains an open issue why affect-related personality traits appear to be less related to the current measures of expertise assessed in non-musicians and non-artists. We also suggest that follow-up studies should include the Big 5 personality traits alongside measures of affect-related traits to gain insights into the intercorrelations between these measures.

More generally, our findings corroborate the need for greater consideration of individual differences in modelling aesthetic experiences, going beyond the investigation of factors related to personality and expertise (Pelowski et al., 2016). Being interested in the immediate subjective experience and not in long-term preferences, we observed significant random effects related to participants in all models, and further, showed that familiarity is not only a significant factor regarding random effects, but that familiarity emerged as a positive predictor in most models for complexity and arousal. We thus suggest that individual differences regarding familiarity should be directly measured and integrated into statistical analysis of data concerning aesthetic experiences. Marin and Leder (2013, 2016) and Marin et al. (2016) already directly assessed familiarity in their experiments using similar stimulus sets and considered familiarity effects in their analyses based on aggregated data across participants. In female university students, familiarity was in most cases negatively correlated with complexity and arousal ratings. However, the mixed-effects modelling technique applied here revealed different findings, which rather suggest a positive association between familiarity and complexity and between familiarity and arousal at the level of the individual. A possible reason for this discrepancy may be grounded in the distinction between correlating average ratings versus averaging individual correlations, which for instance, has been investigated and discussed by Monin and Oppenheimer (2005) in relation to the "beauty is familiar effect". They point out that inter-individual differences are not accounted for and that one needs to be careful in the interpretation of results based on aggregated data.

Following a comparative approach across sensory domains (Marin, 2015), we observed that all measures of individual differences, except for general art interest, predicted either complexity or arousal ratings. This clearly suggests that it is theoretically meaningful to study these affect-related traits in relation to the processing of affective stimuli. Some interesting differences across stimulus categories emerged: trait EI was not a predictor of complexity ratings of visual stimuli but predicted musical complexity. Stress reactivity was a predictor of complexity ratings of cartoons and paintings but not of environmental scenes. Moreover, the arousal models for cartoons and paintings were highly similar. Affective traits similarly predicted both complexity and

arousal ratings of music due to the strong relationship between these two concepts, which was not the case for models of visual stimuli. We thus propose to extend this relatively unexplored avenue of research and to systematically investigate the role of individual differences regarding traits and expertise across different stimulus categories (see also Vessel et al., 2014), such as abstract artworks and patterns, architecture and landscapes, as well as different musical styles.

Although we were able to demonstrate that affect-related traits and expertise may be associated with the processing of complexity and arousal in response to visual and musical stimuli, our study is limited by having concentrated on female university students with generally low to moderate art interest and musical sophistication. Future studies will have to confirm that affect-related traits play less of a role in males, and further, replicate these findings in a group of trained artists and musicians. Moreover, this work, as well as earlier studies by Marin and Leder (2013, 2016, Marin et al., 2016) examined subjective responses to affective visual stimuli only, and it is yet unclear whether the results can be generalized to abstract visual stimuli. In a similar vein, other types of musical styles than piano music need to be studied in the future. Another limitation concerns the exclusive use of self-report measures, which are often susceptible to social desirability, which calls for the future inclusion of ability-based measures, such as ability EI (see Mayer et al., 1990), as well as for additional psychophysiological and endocrine measures of stress.

To conclude, we demonstrated that affect-related traits should be integrated into arousal-based theories of aesthetic experiences, such as the one by Berlyne (1971), and that the impact of these traits on perceived complexity and felt arousal varies across stimulus categories. Our studies may thus lead to a research program that systematically investigates the role of affect-related traits in the processing of other types of affective stimuli, not only in females, but also in men. General art interest and musical sophistication had no impact on arousal responses, but musical sophistication predicted musical complexity ratings, suggesting that measures of expertise and interest may have a differential impact on affective and cognitive processes across stimulus categories. Our studies are thus in line with previous work suggesting that individual differences in aesthetic responses deserve more attention.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2017.12.027.

Funding

This research did not receive any specific grant form funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

We thank Bruno Gingras for helpful comments on an earlier version of the manuscript and Bruno Gingras and Andreas Gartus for their statistical advice.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.
Berlyne, D. E. (1971). *Aesthetics and psychobiology.* New York: Appleton-Century-Crofts.
Berlyne, D. E. (1974). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation.* New York: John Wiley & Sons.
Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition & Emotion, 19*, 1113–1139.
Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion, 24*, 561–595.
Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion, 1*, 300–319.
Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology, 44*, 364–373.
Chmiel, A., & Schubert, E. (2017). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music, 45*(6), 886–909.
Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*, 393–399.
Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression.* New York: Chapman & Hall.
Eerola, T., Vuoskoski, J. K., & Kautiainen, H. (2016). Being moved by unfamiliar sad music is associated with high empathy. *Frontiers in Psychology, 7*, 1176.
Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121–138.
Fayn, K., Maccann, C., Tiliopoulos, N., & Silvia, P. J. (2015). Aesthetic emotions and aesthetic people: Openness predicts sensitivity to novelty in the experiences of interest and pleasure. *Frontiers in Psychology, 6*, 1877.
Filkowski, M. M., Olsen, R. M., Duda, B., Wanger, T. J., & Sabatinelli, D. (2016). Sex differences in emotional perception: Meta analysis of divergent activation. *NeuroImage, 147*, 926–933.
Fox, J., & Weisberg, S. (2010). *An R companion to applied regression.* Thousand Oaks: Sage.
Fox, J., Weisberg, S., Friendly, M., Hong, J., Andersen, R., Firth, D., & Taylor, S. (2016). *Package "effects".* available at The Comprehensive R Archive Network (CRAN). ftp://cran.r-project.org/pub/R/web/packages/effects/effects.pdf.
Freedberg, D., & Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences, 11*, 197–203.
Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception, 28*, 279–296.
Gartus, A., Klemer, N., & Leder, H. (2015). The effects of visual context and individual differences on perception and evaluation of modern art and graffiti art. *Acta Psychologica, 156*, 64–76.
Gartus, A., & Leder, H. (2014). The white cube of the museum versus the gray cube of the street: The role of context in aesthetic evaluations. *Psychology of Aesthetics, Creativity, and the Arts, 8*, 311–320.
Gerger, G., Pelowski, M., & Leder, H. (2017). Empathy, Einfühlung, and aesthetic experience: The effect of emotion contagion on appreciation of representational and abstract art using fEMG and SCR. *Cognitive Processing,* 1–19.
Gingras, B., Marin, M. M., Puig-Waldmüller, E., & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in Human Neuroscience, 9*, 619.
Gökçen, E., Petrides, K. V., Hudry, K., Frederickson, N., & Smillie, L. D. (2014). Subthreshold autism traits: The role of trait emotional intelligence and cognitive flexibility. *British Journal of Psychology, 105*, 187–199.
Greenberg, D. M., Müllensiefen, D., Lamb, M. E., & Rentfrow, P. J. (2015). Personality predicts musical sophistication. *Journal of Research in Personality, 58*, 154–158.
Güçlütürk, Y., Jacobs, R. H.a. H., & Van Lier, R. (2016). Liking versus complexity: Decomposing the inverted U-curve. *Frontiers in Human Neuroscience, 10*, 112.
Hare, F. G. (1974). Artistic training and responses to visual and auditory patterns varying in uncertainty. In D. E. Berlyne (Ed.). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation* (pp. 159–168). Washington: Hemisphere Pub. Corp.
Istók, E., Brattico, E., Jacobsen, T., Krohn, K., Müller, M., & Tervaniemi, M. (2009). Aesthetic responses to music: A questionnaire study. *Musicae Scientiae, 13*, 183–206.
Kokkinos, C. M., & Kipritsi, E. (2012). The relationship between bullying, victimization, trait emotional intelligence, self-efficacy and empathy among preadolescents. *Social Psychology of Education, 15*, 41–58.
Kreutz, G., Schubert, E., & Mitchell, L. A. (2008). Cognitive styles of music listening. *Music Perception, 26*, 57–73.
Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS). Affective ratings of pictures and instruction manual.* Technical report a-7FL: Gainesville.
Marin, M. M. (2015). Crossing boundaries: Toward a general model of neuroaesthetics. *Frontiers in Human Neuroscience, 9*, 443.
Marin, M. M., & Bhattacharya, J. (2013). Getting into the musical zone: Trait emotional intelligence and amount of practice predict flow in pianists. *Frontiers in Psychology, 4*, 853.
Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience, 10*, 536.
Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS One, 8*, e72412.
Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica, 163*, 38–58.
Martins, A., Ramalho, N., & Morin, E. (2010). A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Personality and Individual Differences, 49*, 554–564.
Mayer, J. D., Dipaolo, M., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment, 54*, 772–781.
Mikolajczak, M., Roy, E., Luminet, O., Fillée, C., & de Timary, P. (2007). The moderating impact of emotional intelligence on free cortisol responses to stress. *Psychoneuroendocrinology, 32*, 1000–1012.

Mikutta, C. A., Maissen, G., Altorfer, A., Strik, W., & König, T. (2014). Professional musicians listen differently to music. *Neuroscience, 268*, 102–111.
Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition, 23*, 257–278.
Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. (2014). The musicality of non-musicians: An index for measuring musical sophistication in the general population. *PLoS One, 9*, e89642.
Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts, 28*, 173–191.
Nater, U. M., Abbruzzese, E., Krebs, M., & Ehlert, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology, 62*, 300–308.
Nishikawa, K., & Amemiya, T. (2017). Effects of trait curiosities on the appraisals of picture stimuli: Diversive curiosity and specific curiosity. *International Journal of Affective Engineering, 16*(1), 21–25.
Orr, M. G., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception, 22*, 583–611.
Palmer, S. E., & Griscom, W. S. (2013). Accounting for taste: Individual differences in preference for harmony. *Psychonomic Bulletin & Review, 20*, 453–461.
Papousek, I., Freudenthaler, H. H., & Schulter, G. (2008). The interplay of perceiving and regulating emotions in becoming infected with positive and negative moods. *Personality and Individual Differences, 45*(6), 463–467.
Park, M., Gutyrchik, E., Bao, Y., Zaytseva, Y., Carl, P., Welker, L., ... Meindl, T. (2014). Differences between musicians and non-musicians in neuro-affective processing of sadness and fear expressed in music. *Neuroscience Letters, 566*, 120–124.
Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in Human Neuroscience, 10*, 160.
Petrides, K. V. (2010). Trait emotional intelligence theory. *Industrial and Organizational Psychology, 3*, 136–139.
Petrides, K. V., & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology, 36*, 552–569.
Petrides, K. V., Niven, L., & Mouskounti, T. (2006). The trait emotional intelligence of ballet dancers and musicians. *Psicothema, 18*, 101–107.
Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 21437–21441.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.
Saarni, C. (1999). *The development of emotional competence.* New York: Guilford.
Saklofske, D. H. (1975a). Aesthetic complexity and exploratory behavior. *Perceptual and Motor Skills, 41*, 363–368.
Saklofske, D. H. (1975b). Visual aesthetic complexity, attractiveness and diversive exploration. *Perceptual and Motor Skills, 41*, 813–814.
Schaal, N. K., Bauer, A. K. R., & Muellensiefen, D. (2014). The Gold-MSI: Replication and validation of a questionnaire instrument for the measurement of musical sophistication on the basis of a german sample. *Musicae Scientiae, 18*, 423–447.
Schlotz, W. (2013). Stress reactivity. In M. D. Gellman, & J. R. Turner (Eds.). *Encyclopedia of behavioral medicine* (pp. 1891–1894). New York, NY: Springer New York.
Schmitz, G. M., & von Salisch, M. (2002). Emotionale Selbstwirksamkeit. In: (22.02.2017). http://userpage.fu-berlin.de/gesund/skalen/Emotionale_Selbstwirksamkeit/emotionale_selbstwirksamkeit.htm.
Schott, G. D. (2015). Pictures of pain: Their contribution to the neuroscience of empathy. *Brain, 138*, 812–820.
Schulz, P., Jansen, L. J., & Schlotz, W. (2005). Stressreaktivität: Theoretisches Konzept und Messung [Stress reactivity: theoretical concept and measurement]. *Diagnostica, 51*, 124–133.
Silvia, P. J. (2005). Cognitive appraisals and interest in visual art: Exploring an appraisal theory of aesthetic emotions. *Empirical Studies of the Arts, 23*(2), 119–133.
Silvia, P. J. (2013). Interested experts, confused novices: Art expertise and the knowledge emotions. *Empirical Studies of the Arts, 31*, 107–115.
Silvia, P. J., Fayn, K., Nusbaum, E. C., & Beaty, R. E. (2015). Openness to experience and awe in response to nature and music: Personality and profound aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts, 9*, 376–384.
Singh, Y., & Sharma, R. (2012). Relationship between general intelligence, emotional intelligence, stress levels and stress reactivity. *Annals of Neurosciences, 19*, 107–111.
Solt, F., & Hu, Y. (2015). *Interplot: Plot the Effects of Variables in Interaction Terms.* available at The Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=interplot.
Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia, 50*, 1578–1593.
Street, N., Forsythe, A. M., Reilly, R., Taylor, R., & Helmy, M. S. (2016). A complex story: Universal preference vs. individual differences shaping aesthetic response to fractals patterns. *Frontiers in Human Neuroscience, 10*, 213.
Thoma, M. V., Scholz, U., Ehlert, U., & Nater, U. M. (2012). Listening to music and physiological and psychological functioning: The mediating role of emotion regulation and stress reactivity. *Psychology & Health, 27*, 227–241.
Van der Linden, D., Tsaousis, I., & Petrides, K. V. (2012). Overlap between general factors of personality in the big five, giant three, and trait emotional intelligence. *Personality and Individual Differences, 53*(3), 175–179.
van Paasschen, J., Bacci, F., & Melcher, D. P. (2015). The influence of art expertise and training on emotion and preference ratings for representational and abstract artworks. *PLoS One, 10*(8), e0134241.
Vernon, P. A., Villani, V. C., Schermer, J. A., & Petrides, K. V. (2008). Phenotypic and genetic associations between the big five and trait emotional intelligence. *Twin Research and Human Genetics, 11*, 524–530.
Vessel, E. A., Stahl, J., Maurer, N., Denker, A., & Starr, G. G. (2014). Personalized visual aesthetics. *IS&T/SPIE electronic imaging* (pp. 90140S). International Society for Optics and Photonics.
Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology, 71*, 74.
Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., ... Weil, L. (2006). Development of short forms of the empathy quotient (EQ-short) and the systemizing quotient (SQ-short). *Personality and Individual Differences, 41*, 929–940.
Wiens, S., Sand, A., & Olofsson, J. K. (2011). Nonemotional features suppress early and enhance late emotional electrocortical responses to negative pictures. *Biological Psychology, 86*, 83–89.
Wöllner, C. (2012). Is empathy related to the perception of emotional expression in music? A multimodal time-series analysis. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 214.

2.5.4 Summary and outlook

In Experiment 1, our personality measures were only partly related with each other. In line with the literature, there was a positive correlation between trait emotional intelligence and empathy ($r \sim .50$), and a negative correlation between trait emotional intelligence and stress reactivity ($r \sim -.40$). We were able to demonstrate that among a large group of females affect-related traits were associated with complexity and arousal ratings of visual stimuli, except for complexity ratings of environmental scenes. General art interest was not found to be related to these ratings, which is in line with van Paasschen et al. (2015). However, future studies may try to assess art interest that directly refers to the genre and/or style of the stimuli being shown in the experiments.

To be specific, we did not observe any significant fixed effects for complexity ratings of environmental scenes. Only random effects based on familiarity were significant. For the complexity model of environmental scenes converted into cartoons, stress reactivity emerged as a significant predictor besides the significant random effects. Interestingly, the model for representational paintings was the most complex one. Familiarity was a significant fixed effect and there was an interaction between empathy and stress reactivity. Overall, the perception of visual complexity of environmental scenes seems to be less prone to individual differences than complexity perception of cartoons and paintings (Vessel et al., 2018). One could speculate that this is due to the evolutionary significance of complexity in environmental scenes (Kaplan & Kaplan, 1989). In other words, since complexity is such an important feature of the human environment, humans' survival may decrease if individual differences would interfere to a large extent with basic perceptual processes. It can be argued that the individual differences in preferences for simple vs. complex abstract patterns as described by Güçlütürk et al. (2016) are probably a finding that is specific for this stimulus type, a hypothesis that warrants further testing.

Arousal ratings of all three picture categories were related to familiarity (fixed effect and random effects) and personality traits: for environmental scenes, empathy interacted with stress reactivity and trait emotional intelligence, respectively. For cartoons, there was only one significant interaction, namely between trait emotional intelligence and empathy. This interaction was also significant for representational paintings. Overall, the pattern of results for cartoons and paintings was quite similar and clearly distinct from the one obtained for environmental scenes. Further studies

may include the long form of the trait emotional intelligence questionnaire to gain a better understanding of how empathy interacts with subdimensions of trait emotional intelligence (see also Hoemann et al., 2021, for the manifold concept of emotional expertise). Furthermore, studying contemporary (abstract) visual stimuli (Augustin & Leder, 2006) may reveal to what degree the current results are specific to visual stimuli with clear semantic contents.

In Experiment 2, affect-related traits interacted with complexity and arousal ratings. For complexity ratings, there was a 3-way interaction between stress reactivity, trait emotional intelligence and emotional self-efficacy in addition to two 2-way interactions between stress reactivity and one of the other two measures. A similar pattern of results was present for arousal ratings (note the strong correlation between complexity and arousal in music), but instead of a 3-way interaction, there was an additional interaction between trait emotional intelligence and emotional self-efficacy. Musical sophistication was associated with complexity judgements (but not with arousal): higher musical sophistication was associated with higher complexity ratings (but see Orr & Ohlsson, 2005). As in Experiment 1, familiarity played an important role and emerged as a significant fixed effect and in the significant random effects structure. The finding that stress reactivity interacted significantly with other affect-related traits is in line with Nater et al. (2006), who reported that gender differences in emotional responses to music can be explained by stress reactivity. More recently, Gingras et al. (2015) investigated pupillary responses to Romantic piano trios and related them to ratings of familiarity, arousal, tension and pleasantness. They considered individual differences, such as gender, the role of music in participants' lives and stress reactivity in their analyses. Pupil dilation was predicted by arousal, gender and the role of music and liking for the excerpts, indicating that even psychophysiological responses to music are prone to individual differences of various types. Future studies on music and individual differences should consider both, subjective and psychophysiological responses.

By studying top-down influences on subjective responses to stimuli of different aesthetic quality we could demonstrate that ratings of individual core variables such as complexity and arousal are related to personality traits and expertise, and further, that the stimulus type also determines these relationships. We employed linear mixed-effects modelling, an advanced statistical technique that is not based on averaging across participants' responses but on individual responses, and further, does

not treat individual differences as “error” (see also Pombo & Pelli, 2022, for arguing against a stimulus-based approach in the study of individual differences). Our results regarding familiarity and its association with complexity and arousal demonstrate that there is a discrepancy between correlations based on averaged ratings per stimulus (Marin & Leder, 2013, 2016; Marin et al., 2016) and those based on individual responses without averaging across participants (current study), which has important implications for developing theories of aesthetic experiences (see also Monin & Oppenheimer, 2005).

Individual differences should be taken into account when modelling aesthetic experiences (in the tradition of Berlyne and beyond) already at the stage of considering aesthetic core variables and not only when looking at relationships between these variables (as in Güçlütürk et al., 2016; Güçlütürk & van Lier, 2019). Several current models of aesthetic experiences have not integrated personality into their framework (see Pelowski et al., 2016). Moreover, Weigand and Jacobsen (2023) have recently shown that the intensity of aesthetic experiences may be related to dispositional positive affect. Thus, as a next step, one could examine whether and how the relationship between complexity and arousal is moderated by affect-related traits and expertise across the four stimulus sets. This would reveal how individual differences regarding individual ratings map onto more complex relationships between variables.

The study of individual differences in aesthetic experiences is a timely endeavour (e.g., Chen et al., 2022; Hogan, 2015; Pombo & Pelli, 2022). On the one hand, there are regularities and general tendencies to be observed when studying aesthetic experiences or preferences more generally. On the other hand, individual differences (some can be explained by mood, expertise, age, gender and personality) have been documented. For instance, the study by Vessel et al. (2018) compared shared and individual aesthetic preferences for natural landscapes, exterior architecture, faces and paintings. They found that for natural landscapes and faces shared taste was stronger, whereas for architecture and paintings preferences varied more across participants. Vessel et al. (2018) also pointed out that stronger shared taste does not mean that people rely more on stimulus features (for similar studies see Brielmann & Pelli, 2019; Leder et al., 2016). These findings are partly in line with our results, in which we observed that complexity ratings of environmental scenes were not related to personality traits, but personality was associated with complexity perception when viewing cartoons (stress reactivity), and to an even stronger degree,

when viewing paintings (stress reactivity and empathy). Arousal ratings of visual scenes were generally more affected by personality traits than complexity ratings, which could be explained by the fact that arousal refers to an inner subjective state and complexity to the perception (and description) of a stimulus of the outside world. Subjective self-reports of inner states and preferences may be more influenced by individual differences than ratings that refer to stimulus characteristics. This hypothesis could be explored in greater depth in the future.

2.6 Gaze patterns reveal aesthetic distance while viewing art (Marin & Leder, 2022)

2.6.1 Introduction

Much of the current debate over the nature of aesthetic experiences has revolved around the neurobiological basis of aesthetic experiences (e.g., Chatterjee & Cardillo, 2022). In this study, we were interested in employing a different methodology, namely eye tracking, to study characteristics of visual aesthetic experiences by comparing exploratory behaviour across environmental scenes and representational paintings. Eye-tracking methodology allows for a free-viewing task, which does not require additional subjective ratings. Thus, this approach enhances the ecological validity of research results in the study of aesthetic experiences because it mirrors natural viewing behaviour.

For both experiments included in this study, we selected stimuli from Marin and Leder (2013) that varied in arousal, pleasantness and complexity. As in Marin and Leder (2016), we were also interested in the effect of presentation duration and thus presented our stimuli (picture category as a within-subject factor) for either 25 or 6 s to two groups of participants (Experiments 1 and 2). We matched our two samples on several background measures, including mood prior to the experiment and art interest. Objective measures of complexity (Marin & Leder, 2013), and other image statistics such as spatial frequency, were used to demonstrate that our effects due to subjective complexity, pleasantness and arousal are independent of low-level image features.

We framed the current study within the context of philosophical theories about the distinctive features of aesthetic experiences, such as disinterested pleasure and aesthetic distance (see Cupchik, 2002; Hayn-Leichsenring & Chatterjee, 2019). We predicted that aesthetic experiences can already be identified by the way people behold a painting versus a motivational environmental scene.

2.6.2 Authorship contribution statement

MMM. Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing − original draft preparation, visualization, and project administration. HL. Supervision, resources, writing − reviewing and editing.

2.6.3 Re-print of the published research article

DOI: 10.1111/nyas.14792

ORIGINAL ARTICLE

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Gaze patterns reveal aesthetic distance while viewing art

Manuela M. Marin[1,2] | Helmut Leder[1]

[1]Department of Cognition, Emotion and Methods in Psychology, University of Vienna, Vienna, Austria

[2]Department of Psychology, University of Innsbruck, Innsbruck, Austria

Correspondence
Manuela M. Marin, Department of Cognition, Emotion and Methods in Psychology, University of Vienna, 1010 Vienna, Austria.
Email: manuela.marin@uibk.ac.at

Abstract

For centuries, Western philosophers have argued that aesthetic experiences differ from common, everyday pleasing sensations, and further, that mental states, such as disinterested contemplation and aesthetic distance, underlie these complex experiences. We empirically tested whether basic perceptual processes of information intake reveal evidence for aesthetic distance, specifically toward visual art. We conducted two eye tracking experiments using appropriately matched visual stimuli (environmental scenes and representational paintings) with 59 participants using two different presentation durations (25 and 6 s). Linear mixed-effects models considering individual differences showed that affective content (pleasantness and arousal), but not stimulus composition (complexity), leads to differential effects when viewing representational paintings in comparison to environmental scenes. We demonstrate that an increase in aesthetic pleasantness induced by representational paintings during a free-viewing task leads to a slower and deeper processing mode than when viewing environmental scenes of motivational relevance, for which we observed the opposite effect. In addition, long presentation durations led to an increase in scanning behavior during visual art perception. These empirical findings inform the debate about how aesthetic experiences differ from everyday perceptual processes by showing that the notion of aesthetic distance may be better understood by examining different modes of viewing.

KEYWORDS
disinterested contemplation, emotional distance, empirical aesthetics, eye movements, scene perception

INTRODUCTION

Humans are uniquely characterized by their ability to experience pleasure from perceiving either human-made art (e.g., paintings and music) or other objects and events which are part of the natural world (e.g., animals, faces, and landscapes).[1–4] How humans respond aesthetically and make aesthetic judgments has been a central subject of theoretical speculation in philosophical discourse since Antiquity. According to philosophical treatises, during aesthetic episodes, these natural or human-made objects inducing pleasure are of little motivational relevance, do not require an immediate behavioral response, and are thus unrelated to immediate survival.[5–7] However, from a biological and psychological perspective, still little is known about what distinguishes pleasurable aesthetic encounters from the perception of more mundane objects of motivational relevance.[8]

Art has been defined as the process of turning an object or activity into something special, being able to generate complex aesthetic experiences in humans.[9] In Western societies, aesthetic experiences

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of New York Academy of Sciences.

have been mostly decoupled from lower, sensuous pleasures and are often seen to result from disinterested contemplation, as proposed by 18th-century philosophers, such as Shaftesbury, Hutcheson, and Kant. Similar ideas were later propagated by Schopenhauer, Beardsley, Stolnitz, and other modern philosophers. In Kantian philosophy, judgments on the beautiful and on the sublime are grounded in specific feelings of pleasure, and this pleasure is thought to be of a distinctive, disinterested kind. For example, a disinterested attitude refers to viewing a painting contemplatively by putting its perceptual properties and meaning into the foreground.[10] This stands in contrast to the more general attitude of viewing an object (whether natural or human-made) as something of personal interest in the sense of its usefulness for the viewer, thus referring to utilitarian goals, possible actions, and personal investment.

A closely related concept is described by aesthetic attitude theories, which argue that it is people's aesthetic attitude toward an object that turns it into an aesthetic object. For instance, at the beginning of the 20th century, Bullough, proposing the concepts of *aesthetic consciousness* and *psychical distance*, described aesthetic appreciation as a mental state in which the practical role of an object and personal concerns are disregarded. Aesthetic distance is characterized by a personal relation to the perceived object and can vary in degree, depending on the nature of the object and on the perceiver's capacity to maintain such a state.[5] The concepts of disinterestedness and aesthetic distance have both been subject to theoretical criticism,[11] which may also be related to the fact that it is a challenging task to develop experimental paradigms to demonstrate this unique mode of perception among humans. Importantly, the concepts of aesthetic attitude and aesthetic experience are not identical. The former refers mostly to a special type of attention and perception of an object (i.e., a different perceptual mode in the case of visual aesthetic perception), whereas the latter also involves higher-order cognitive and affective processes leading to aesthetic judgments (e.g., beauty and liking), complex aesthetic emotions (e.g., the sublime and nostalgia), as well as comprehension (e.g., meaning of a painting).[4,12] As a fruitful avenue of research, understanding the nature of aesthetic attitude requires systematic comparisons of perceptual and physiological processes in response to different stimulus categories, within and across sensory domains.[13]

Currently, there are only a few empirical findings that might substantiate the existence of such distanced mental states during art perception. In the context of emotion research, and referring to Kant's idea of disinterestedness, it has been argued that emotional distancing is typical of aesthetic emotions, which differentiates them from utilitarian emotions.[7] By manipulating the context ("This is an artwork" vs. "This is a press photograph") in which pleasant and unpleasant affective environmental scenes and artworks were perceived, researchers were able to distinguish between emotional and aesthetic responses using facial electromyography.[14] In an art context, responses to unpleasant artworks were associated with an increase in activity of the zygomaticus major (smiling muscle), and these artworks were also liked more than in a nonart context. In other words, the art context made people respond less strongly to the negative valence of artworks, suggesting that motivational appraisal may be reduced under this condition. The context had little effect on the processing of environmental scenes.[8] However, in another study, the presentation of the same pleasant and unpleasant environmental scenes presented in art versus nonart contexts led to differences in electrophysiological responses in the brain.[15] Both pleasant and unpleasant pictures had higher liking scores in the art condition than in the nonart condition and were associated with an attenuated late-positive potential (LPP). This result can be interpreted as indicative of implicit emotion regulation because the LPP is modulated by the emotional intensity of a stimulus and related to attentional resources.[16] These studies, although informative about emotional processes, can be criticized for having actively evoked association of art-related perceptions and behaviors in viewers by employing an explicit framing manipulation and asking for subjective ratings while measuring physiological responses, which does not mirror natural viewing behavior (a neuroimaging study by Cupchik *et al.*[17] used a similar methodology). Nevertheless, previous neuroimaging studies have provided some evidence for spontaneous aesthetic processing of stimuli.[18,19]

It is a challenging task to develop experimental paradigms with the goal of revealing a unique mode of perception, such as disinterested contemplation and aesthetic distance, ultimately leading to aesthetic experiences. Because an aesthetic experience is assumed to be extended in time and might be marked by a physiological state corresponding with aesthetic distance, we need to employ a continuous measure that informs us about the possibility of such a special—aesthetic—mental state over a longer time. The analysis of eye movements, as a reliable and established indicator of sensory information processing,[20] can be key in reaching this goal. Eye movement parameters have been found to be task-dependent, particularly when analyzed over the course of a trial.[21–23] For instance, it was recently reported that when participants were asked to make aesthetic judgments of indoor and outdoor scenes (without the depiction of humans), the fixation duration steadily increased over the first 30 fixations, but not when the same everyday images had to be memorized.[21] However, an explicit rating task might hinder a truly aesthetic, distanced experience. Eye tracking also enables free-viewing tasks, which is ideal for measuring natural viewing behavior of images of different qualities.

Since eye movements are influenced by several visual features (e.g., complexity/composition, semantic, and affective contents[20,24]), it is necessary to consider these factors when comparisons across stimulus categories are made. Previous findings on the role of pleasantness and arousal on eye movements have been inconsistent.[25,26] Affective cues of visual scenes engage the appetitive or defensive motivational systems in the brain,[20] therefore, we decided to use pictures of similar semantic contents that vary in pleasantness (unpleasant vs. pleasant) and arousal (low vs. high)[24] alongside complexity. To provide convincing evidence for art-specific, aesthetic differences that can be associated with emotional or aesthetic distance, we need to show that affective content of paintings influences eye movement parameters differently than when viewing environmental scenes of the same affective content. To be specific, if we compare eye movements, ideally within-subjects in relation to stimuli that vary mostly in their motivational relevance and aesthetic nature, such as environmental scenes versus representational paintings (in our study, *picture*

category), differences in scanning behavior should be primarily observed along the pleasantness dimension, which has been associated with motivational salience in empirical studies[20,26] and aesthetic distance in the philosophical literature.[5,10] An aesthetic stance may be reflected by a deeper, more enduring exploration of the more aesthetically pleasing stimuli, with fewer fixations and longer fixation durations, as there is no need to process information quickly to elicit behavioral responses that may facilitate survival. This would be in line with the philosophical assumptions of disinterested pleasure and aesthetic distance. Regarding the effects of arousal and complexity, we predicted that both factors should lead to independent effects[20,24] on eye movement parameters in addition to effects due to pleasantness. However, we refrained from formulating specific hypotheses about possible differential effects when viewing environmental scenes versus representational paintings.

Another issue to be considered concerns the possible association between viewing mode and presentation duration.[27] Since aesthetic experiences are presumed to take some time to unfold,[28,29] images were presented for a duration sufficient to allow aesthetic experiences to take place, that is, a presentation duration of 25 s in experiment 1. In experiment 2, we used a shorter presentation duration of 6 s, which has been commonly employed when studying environmental scenes and eye movements.[20] By comparing presentation durations of 1, 5, and 25 s and using the same stimulus material as in the current study, we have previously shown that subjective experiences of pleasantness, arousal, and complexity may depend on presentation duration.[30] Last, aesthetic distance regarding the perception of art should be an inherent characteristic of any perceiver, so we ensured that art expertise was not particularly high among viewers by administering an art interest questionnaire.

MATERIALS AND METHODS

Stimuli

We selected 32 environmental scenes (pictures from the *International Affective Picture System*[31]) and 32 representational paintings from 192 stimuli used in a previous study.[29,32] The two picture sets were carefully chosen based on their affective and compositional contents, to ensure that they vary in a similar degree in pleasantness (very unpleasant vs. very pleasant), arousal (very calm vs. very excited), as well as visual complexity (very simple vs. very complex). To be specific, the stimulus selection was based on the two-dimensional circumplex model of affect proposed by Russell,[33] and we aimed for four categories of stimuli (low-arousing pleasant, high-arousing pleasant, low-arousing unpleasant, and high-arousing unpleasant) that varied maximally in arousal and pleasantness. Visual complexity varied to an equal degree within the four categories. The comparison between environmental scenes and representational paintings and the consideration of arousal as a factor alongside pleasantness and complexity is an extension of Bradley *et al.*'s eye movement study.[20] Environmental scenes and representational paintings depicted humans in different social settings as well as animals, food, and landscapes. The semantic contents of the pictures did not comprise erotic couples, mutilated/burned bodies, and suffering animals. The affective space covered by the selected stimuli is representative of the affective spaces of the full stimulus sets as in Ref. 29, generally following a U-shaped distribution.

For experiment 1, the stimuli were selected from results of a previous study[29] on the basis of ratings obtained for a presentation duration of 25 s. Thirty-six participants (18 males, mean age 23.9 ± 4.9 years, age range 21–42 years; 18 females, mean age 22.6 ± 4.7 years, age range 20–41 years) rated environmental scenes, and 40 participants (20 males, mean age 24.4 ± 4.1 years; age range 19–38 years; 20 females, mean age 22.0 ± 1.5 years, age range 20–26 years) rated representational paintings on felt pleasantness, arousal, and perceived complexity on 7-point scales. The stimuli were selected based on average ratings computed for each picture. In general, inter-rater reliability was very high for all types of ratings (see Refs. 29 and 32 for further details), which allowed us to interpret the ratings of both picture sets as normative ratings.

For experiment 2, the presentation duration was shortened to 6 s to extend our findings from experiment 1 to shorter aesthetic experiences. To obtain subjective ratings of the 32 environmental scenes and 32 representational paintings selected for experiment 1, they were presented together with all pictures of the respective full stimulus set as described in Ref. 29. This procedure ensured that the context of the subjective ratings was the same for both experiments, although the presentation duration differed considerably. The rating procedure was identical as to that used in experiment 1 (the data of this rating study have not been previously published). Participants were 39 university students (20 females, 20.2 ± 1.1 years, age range 18–22 years; 19 males, 22.5 ± 2.3 years, age range 20–28 years) who rated the environmental scenes, and 43 participants (21 females, 21.0 ± 2.4 years, age range 18–28 years; 22 males, 22.1 ± 1.4 years, age range 20–25 years) who rated the representational paintings. Note that the subjective ratings of pleasantness, arousal, and complexity used for selecting the stimuli for both experiments were also used as fixed factors in the main analyses (linear mixed-effects models [LMMs]).

Tables S1 and S2 show the mean ratings of pleasantness, arousal, and complexity for each picture presented for 25 and 6 s, given for environmental scenes (Table S1) and representational paintings (Table S2), respectively. Table S3 shows the average ratings of pleasantness, arousal, and complexity for each picture category, and Table S4 depicts these ratings summarized according to each affective category varying in pleasantness and arousal. The two stimulus sets are comparable regarding affective contents and perceived visual complexity.

For experiments 1 and 2, the two short versions of the German Multidimensional Mood State Questionnaire (MBDF)[34] were administered before and after the recording of eye movements. This questionnaire assesses positive/negative mood, alertness/fatigue, and quietude/disquietude on 5-point scales. Only the mood assessed prior to the experiment (version A) was considered in the analysis. An earlier version of the Vienna Art Interest and Art Knowledge Questionnaire (VAIAK)[35] was used to assess art interest with 12 items (11 items corresponding to current items 1–11 of the

VAIAK and an additional item measuring the frequency of art exhibitions visited and art knowledge, which was not evaluated for this study).

A short self-developed questionnaire asked questions on the sociodemographic background of the participants, their general liking of the environmental scenes and paintings, from (1) "not at all" to (7) "very much," whether participants generally experienced feelings in response to these pictures, from (1) "never" to (7) "always," and how strong these feelings were, from (1) "very weak" to (7) "very strong." We also assessed the general familiarity with the paintings on a similar 7-point scale.

Participants

Forty-three university students were tested for experiment 1 (long presentation duration, 25 s), and 36 participants for experiment 2 (short presentation duration, 6 s). A total of 14 participants had to be excluded in experiment 1, and six in experiment 2. Participants only took part in one of the two experiments. The exclusion criteria regarding the eye movement measurements comprised technical problems (e.g., issues with calibration due to wearing thick glasses, five cases in experiment 1, one case in experiment 2) and, in one case in experiment 1, extreme fatigue during the eye movement measurement as shown by upper eyelids coming down regularly across several trials. We specifically sought to obtain a similar calibration quality across participants and conditions. Three more participants had to be excluded in experiment 1 and two in experiment 2 after inspection of the eye movement data (for details, see Analysis of eye movement data). Further exclusions were made due to issues with the experimental protocol (e.g., unfilled questionnaires; three in experiment 1 and two in experiment 2) or background variables (e.g., high familiarity with the pictures; two in experiment 1 and one in experiment 2).

The final sample of experiment 1 (sample 1) comprised 29 participants (15 females, 24.5 ± 2.4 years, age range 21–28 years; 14 males, 21.9 ± 1.5 years, age range 20–24 years), and 30 participants (sample 2, 14 females, 21.5 ± 2.0 years, age range 19–25 years; 16 males, 21.5 ± 1.9 years, age range 19–26 years) were included in experiment 2. Sample 1, M = 23.2 years, SD = 2.37, was older than sample 2, M = 21.5 years, SD = 1.91, U = 251.00, z = −2.82, p = 0.005. Sample 1, M = 65.93, SD = 14.78, and sample 2, M = 68.73, SD = 16.5, reported a similar interest in art, U = 386.00, z = −0.74, p = 0.46. There were no significant differences regarding mood states prior to the experiment, $U_{\text{good-bad}}$ = 396.00, z = −0.60, p = 0.55; $U_{\text{awake-tired}}$ = 404.00, z = −0.48, p = 0.64; $U_{\text{calm-nervous}}$ = 388.00, z = −0.72, p = 0.47.

The two samples did not differ regarding their familiarity with the paintings, U = 376.00, z = −0.95, p = 0.35, which was generally low (M_{sample1} = 2.03, SD = 1.15; M_{sample2} = 2.33, SD = 1.27). Similarly, our samples did not differ regarding their general liking of the environmental scenes, U = 360.50, z = −1.16, p = 0.24, M_{sample1} = 4.31, SD = 1.31; M_{sample2} = 4.67, SD = 1.32, and representational paintings, U = 409.50, z = −0.40, p = 0.69, M_{sample1} = 4.93, SD = 1.22; M_{sample2} = 5.00, SD = 1.01.

The two samples reported differences regarding the feelings elicited when viewing environmental scenes, U = 309.00, z = −1.98, p = 0.048, and the degree of these feelings, U = 318.50, z = −1.84, p = 0.066. When environmental scenes were presented for 25 s (experiment 1), feelings were less frequent, M_{feelings} = 4.62, SD = 1.54, and less intense, $M_{\text{intensity}}$ = 3.86, SD = 1.75, than when the scenes were presented for 6 s (experiment 2), M_{feelings} = 5.33, SD = 1.21; $M_{\text{intensity}}$ = 4.60, SD = 1.10. These differences were not observed for feelings induced by paintings. Neither the reported frequency, U = 394.5, z = −0.64, p = 0.52, nor the intensity, U = 411.00, z = −0.37, p = 0.71, differed between the two samples ($M_{\text{feelings1}}$ = 4.90, SD = 1.40; $M_{\text{intensity1}}$ = 3.86, SD = 1.53, $M_{\text{feelings2}}$ = 4.73, SD = 1.23; $M_{\text{intensity2}}$ = 3.77, SD = 1.25).

Procedure

The data were collected following the local guidelines of the Faculty of Psychology at the University of Vienna, Austria, and in accordance with the Declaration of Helsinki (revised 1983). According to the Austrian Universities Act 2002 (UG2002), which held at the time the study was carried out, only medical universities were required to appoint ethics committees for clinical tests, application of medical methods, and applied medical research. Therefore, no ethical approval was required for this study. Written informed consent was given by all participants, who took part voluntarily in the study and could withdraw at any time during the experiment without further consequences.

The procedure was the same for experiments 1 and 2. Participants were first asked to read and sign an informed consent form and received course credits for participating. We then tested the participants' visual acuity (near vision reading test, Oculus 47211), color vision (Ishihara color test), and ocular dominance with a dominant eye test card. Only participants with normal or corrected vision were asked to proceed. The short version A of the MBDF was administered (paper and pencil version). Next, in a quiet, dimly lit room (ambient light levels of 35 lux as measured just below the forehead support using an X-Rite i1 Pro lux meter), participants were asked to position themselves comfortably in front of the computer screen with their heads fixated with a chin and forehead rest (see Supplementary Materials for details on the eye tracking apparatus). To keep the chin and forehead rest at the same height across participants, they were asked to adjust the height of the armchair. Prior to the experiment, female participants were asked to arrive without heavy make-up and mascara. A make-up remover was made available for all participants if required, as well as eye drops for persons wearing contact lenses. Participants read the instructions on the screen and were asked not to move their head during the measurements, after which a 9-point calibration/validation followed. Participants wearing glasses were included if the calibration was acceptable.

The pictures were presented in two blocks grouped according to the picture type (environmental scenes vs. representational paintings) and shown at a size of 34.0 × 25.5 cm on a black background. The order of the two blocks was balanced across participants, and the pictures within each block were presented randomly. In each block,

participants were first shown four pictures as practice trials to familiarize themselves with the procedure before the presentation of the actual 32 pictures began. Participants could take a short break between the two blocks if desired. The interstimulus interval was 5 s, followed by a drift check prior to the next picture onset; therefore, all participants began their scene viewing at the center of each picture. Recalibration was applied whenever the drift check failed. The experimenter carefully monitored the eye movement measurement on a second screen hidden behind a board, ensuring natural viewing behavior of the participants. After the two experimental blocks, participants filled in the short version B of the MBDF, followed by the self-developed one-page questionnaire and the art interest and art knowledge questionnaire (VAIAK). The experimental session lasted approximately 75 (experiment 1) or 50 min (experiment 2), after which participants were thanked and debriefed.

Analysis of eye movement data

The eye movement data were analyzed with Data Viewer 2.6.1 (SR Research). The fixation duration threshold was set to 100 ms, and the saccade amplitude threshold to 0.5°.[36] All data files and trials were carefully screened to ensure that participants looked at the picture during the complete presentation duration and that the quality of the eye movement data was of a high standard. First, an interest area report, based on the picture size on the screen, was generated per trial to obtain a percentage of fixations that were found on the picture for a given presentation duration (either 25 or 6 s). Values above 95% were considered as acceptable, and values below 80% as insufficient. All critical trials lying between 80% and 95% were checked individually (by author M. Marin). For example, it was possible that some fixations were found at the edge of a picture because an object was shown close to the edge. Such trials were included in the data. In general, the quality of the remaining eye movement data (29 participants for experiment 1 and 30 participants for experiment 2) was very high, and only a few trials were removed (coded as NA) from the individual participants' data used in the main analyses (1.5% of the 938 trials in the IAPS 25-s condition and 0.9% of the 928 trials in the paintings 25-s condition, 0.83% of the 960 trials in the IAPS 6-s condition and 0% in the paintings 6-s condition).

Analysis of image statistics

As in the study of Bradley *et al.*,[20] we also used image statistics to determine physical differences between images and to test whether emotion modulates eye movements when other potential factors are considered in the statistical LMMs. For this purpose, we computed nine measures: first, objective measures of complexity based on JPEG compression file size, edge detection (perimeter detection and Canny algorithm), entropy, RMS contrast-alpha, and phase congruency-beta (see the study by Marin and Leder[29] for further details on the respective analyses). Second, we computed the average brightness of an image in the RGB (red, green, blue) color space in Matlab (Mathworks Inc., Sherborn, MA, USA). Third, spatial frequencies were determined by a discrete wavelet analysis.[37] The frequency bands analyzed with the wavelet in Matlab[37] were grouped into two measures: low-spatial frequency (two bands: gray 4 and gray 8) and high-spatial frequency (five bands: gray 32 to gray 512). Intercorrelations between these nine objective measures for environmental scenes (Table S19) and for representational paintings (Table S20) reveal only two incidences of multicollinearity (r_s = >0.80). In both picture sets, high spatial frequency correlated highly with the RMS contrast alpha measure. Thus, the RMS contrast measure was discarded from further analyses.

Statistical analysis

A series of LMMs were computed using R 3.4.2 (see https://www.r-project.org) to predict eye movement parameters (fixation count, fixation duration, and saccade amplitude). LMMs are powerful statistical tools because they do not require aggregating data over pictures and participants, and they allow for the modeling of individual differences by using random intercepts and slopes. Random error is modeled at all levels simultaneously by using maximum-likelihood procedures to estimate coefficients.[38] Both between- and within-subject effects are accounted for in the analysis,[39] and models can lead to broad generalizations.[40] The *lmer* function in R package lme4 was used to compute the LMMs,[38] and significance tests were calculated using the *anova* function from the package car.[41] To determine influential cases among participants, Cook's distances[42] were calculated for a full model that included all main effects and interactions of interest (influence package). More details about influential cases and other aspects of the analysis can be found in the Supplementary Materials.

The random effects structure was determined for each individual model. A random intercept model, consisting of by-subject random intercepts and by-stimulus random intercepts, was compared to the null-model using a likelihood ratio test.[43] For each model, the results indicated that the random intercept models were a significant better fit than the respective null models. All final models were random-intercept models. Pictures and participants were treated as fully crossed random effects.[44]

For each experiment, we first computed three LMMs (one model for each eye movement parameter), which directly compared the scanning behavior when viewing environmental scenes versus representational paintings (main analysis). Since we observed significant interactions between picture category and pleasantness in both experiments, we then computed, for each presentation duration, three models for each picture category separately, and again one model for each eye movement parameter (see Supplementary Materials). Last, we considered objective measures related to image statistics as predictors and computed three additional models for each picture category in both experiments (see main analysis and Supplementary Materials).

In our main analysis (Figures 1 and 2), participants, pictures, and picture category were coded as categorical factors (1 = representational paintings and 2 = environmental scenes), whereas all other predictors

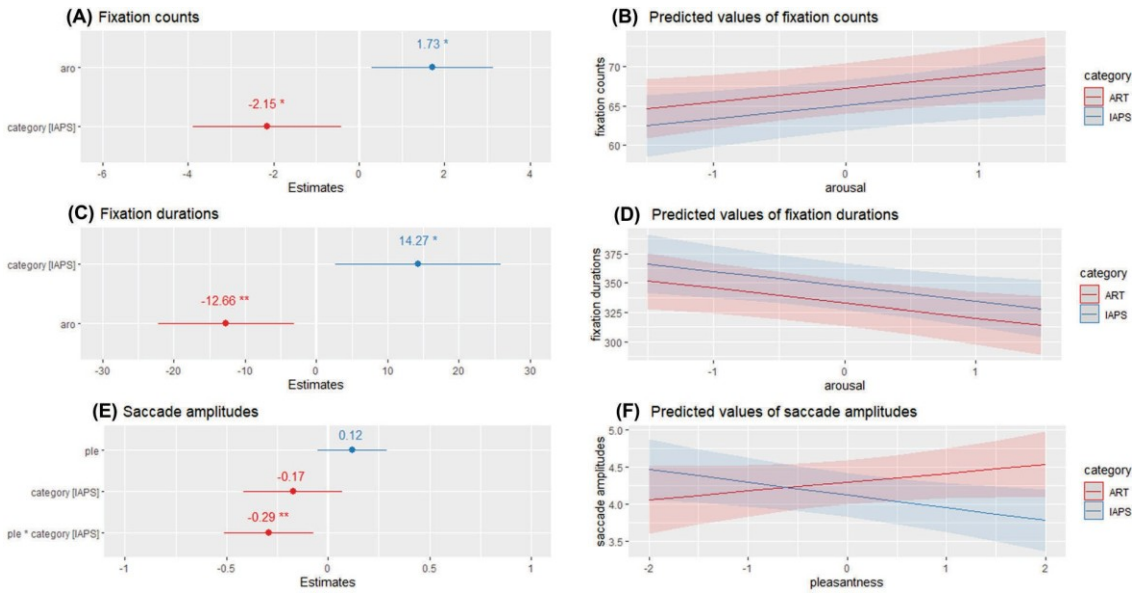

FIGURE 1 Estimates and marginal effects of three linear mixed-effects models comparing eye movements across environmental scenes and representational paintings presented for 25 s. (A) Coefficient estimates of linear mixed-effects model predicting average fixation counts. (B) Marginal effects of arousal and picture category predicting fixation counts. The effect of picture category was significant, with a higher number of fixations when viewing paintings compared to environmental scenes. An increase in arousal led to an increase in fixation counts for both picture types. In the model for fixation counts, an interaction between pleasantness and picture category was present at trend level. (C) Coefficient estimates of linear mixed-effects model predicting average fixation durations. An increase in arousal was associated with shorter presentation durations, and fixation durations were shorter when viewing art than environmental scenes. (D) Marginal effects of arousal and picture category predicting fixation durations. (E) Coefficient estimates of linear mixed-effects model predicting saccade amplitudes. There was a significant pleasantness × category interaction. A three-way arousal × complexity × pleasantness interaction was present at trend level. Note that the main effect of pleasantness refers to the arts category and does not indicate a general main effect in this case. (F) Marginal effects of the interaction between pleasantness and picture category predicting saccade amplitudes. Error bars/bands indicate 95% confidence intervals. In all models, fixed factors included arousal (aro), pleasantness (ple), complexity (com), and picture category (ART vs. IAPS), as well as all possible interaction terms between these factors. ART = representational paintings (reference category), IAPS = environmental scenes. Both models are based on 1834 observations, 64 pictures, and 29 participants. Random intercepts for pictures and participants were highly significant for all models. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$

were treated as continuous covariates and grand mean centered.[45] The set of predictors (fixed factors) included complexity, pleasantness, and arousal ratings for each picture, as well as the respective two-way and three-way interaction terms. When included as predictors in the models, objective measures related to image statistics were *z*-transformed. For both stimulus sets, the fixed factors included JPEG file size, perimeter detection, Canny algorithm, phase congruency-beta, entropy, brightness, low spatial frequency, high spatial frequency, as well as complexity, pleasantness, and arousal ratings (no interaction terms were included).

RESULTS

Since we employed a within-subject design in each experiment, it was possible to directly compare eye movements when viewing environmental scenes and representational paintings. In a first step, LMMs were fitted to the eye movement data to model random intercepts (for both pictures and participants), as well as fixed factors of pleasantness, arousal, complexity, and picture category. Second, we analyzed the data separately for each picture category and presentation duration by presenting results for environmental scenes first. Finally, we present a more complete analysis involving objective measures related to image statistics to examine whether emotional effects on eye movements persist when other potential modulators are removed.

In both experiments (Figures 1 and 2), models predicting either fixation counts, fixation durations, or saccade amplitudes revealed significant interactions between pleasantness and picture category. In experiment 1 (see Figure 1A–F for details), we observed significant main effects of picture category when modeling fixation counts (Figure 1A) and fixation durations (Figure 1C). Mean fixation counts were higher for paintings (M = 67.08, SD = 11.23) than for environmental scenes (M = 65.25, SD = 11.95). Mean fixation durations were shorter for paintings (M = 334.09 ms, SD = 73.23) than for environmental scenes (M = 345.87 ms, SD = 100.22). There was also a main effect of arousal on these measures (Figure 1A,C), indicating that an increase in arousal was associated with a higher number of fixations and shorter fixation durations. For saccade amplitudes (Figure 1E), a significant

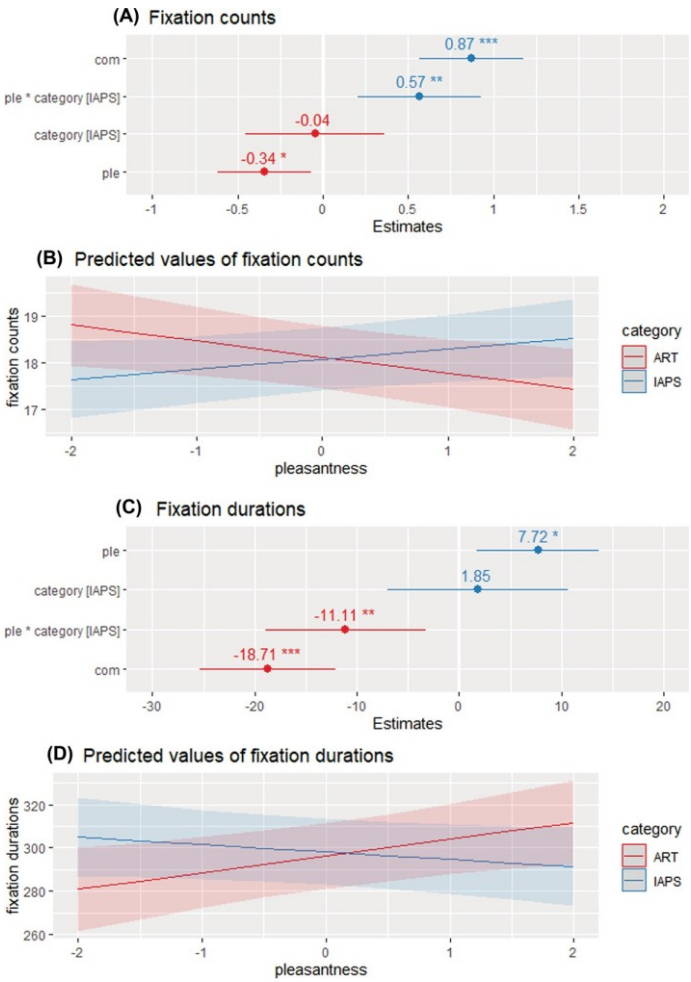

FIGURE 2 Estimates and marginal effects of two linear mixed-effects models comparing eye movements across environmental scenes and representational paintings presented for 6 s. (A) Coefficient estimates of linear mixed-effects model predicting average fixation counts. The effect of picture category was not significant, and an increase in complexity led to an increase in fixation counts (and a decrease in fixation durations) for both picture types. Note that the main effect of pleasantness (also in C) only refers to the arts category and is not a general main effect. (B) Marginal effects of two-way interaction between pleasantness and picture category. An interaction between pleasantness and picture category emerged, indicating that an increase in pleasantness led to a reduction in scanning behavior (fewer fixations and longer fixation durations) when viewing visual art. (C) Coefficient estimates of linear mixed-effects model predicting average fixation durations. (D) Marginal effects of two-way interaction between pleasantness and picture category. Error bars/bands indicate 95% confidence intervals. In both models, fixed factors included arousal (aro), pleasantness (ple), complexity (com), and picture category (ART vs. IAPS), as well as all possible interaction terms between these factors. ART = representational paintings (reference category), IAPS = environmental scenes. Both models are based on 1848 observations, 64 pictures, and 29 participants (one influential case was removed). Random intercepts for pictures and participants were highly significant for all models. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$

pleasantness × category interaction emerged. The more pleasant a painting was, the longer were the saccade amplitudes. In experiment 2 (Figure 2A–D), backward elimination revealed fixed effects of complexity and pleasantness, as well as an interaction between pleasantness and picture category in both models predicting average fixation counts (Figure 2A) and fixation durations (Figure 2C). More complex pictures were associated with a higher number of fixation counts and shorter fixation durations, regardless of the picture category. Pleasant paintings, in comparison to environmental scenes, were associated with fewer fixation counts (Figure 2B) and longer fixation durations (Figure 2D). The model predicting average saccade amplitudes (1848 observations, 64 pictures, and 29 participants, not shown in Figure 2) revealed no significant fixed effects (random effects: 1 | participants, $\chi^2 = 458.79$, p(LRT) < 0.001; 1 | pictures, $\chi^2 = 574.74$, p(LRT) < 0.001).

Due to the observed significant pleasantness × picture category interactions in both experiments (and two other interactions at the trend level in experiment 1), we also provide separate follow-up analyses for environmental scenes and representational paintings (see Supplementary Materials and Tables S5–S16). In experiment 1, more complex ($B(SE)$ = 2.81(0.41), 95% CI = [2.03, 3.60]) and pleasant ($B(SE)$ = 0.97(0.32), 95% CI = [0.35, 1.59]) environmental scenes led to a larger number of fixations (i.e., shorter fixation durations, Tables S5 and S6), with an additional positive effect of emotional arousal ($B(SE)$ = 1.46(0.60), 95% CI = [0.30, 2.62]). Moreover, the more pleasant a picture was, the shorter was the saccade amplitude ($B(SE)$ = −0.21(0.07), 95% CI = [−0.35, −0.07], Table S7). A similar pattern was observed for environmental scenes presented for 6 s in experiment 2 (Tables S8–S10). In summary, these results based on long and short presentation durations suggest an increase in information-seeking and scanning behavior due to affective and compositional image features when viewing environmental scenes.

For representational paintings presented for 25 s (experiment 1), more pleasant pictures led to a smaller number of fixations ($B(SE)$ = −1.31(0.52), 95% CI = [−2.35, −0.27]) and longer fixation durations ($B(SE)$ = 8.25(3.00), 95% CI = [2.29, 14.21]) (Tables S11 and S12). There were no effects of complexity and pleasantness on saccade amplitudes (Table S13) and thus spatial explorative behavior. For the shorter presentation duration of 6 s (experiment 2, Tables S14–S16), the effect of pleasantness on fixation counts was similar ($B(SE)$ = −0.35(0.15), 95% CI = [0.03, 1.13]), and there was

TABLE 1 Results of linear mixed-effects model with subjective and objective fixed factors predicting average fixation counts when viewing environmental scenes for 25 s

Fixed effects	*B (SE)*	95% CI	*df*	χ^2	*p(LRT)*
(Intercept)	65.23 (1.55)	[62.16, 68.30]			
Perimeter detection	−1.05 (0.46)	[−1.91, −0.19]	1	5.55	0.018*
Complexity	3.59 (0.49)	[2.67, 4.50]	1	34.51	<0.001***
Pleasantness	1.15 (0.32)	[0.54, 1.76]	1	11.91	<0.001***
Arousal	1.63 (0.61)	[0.49, 2.78]	1	7.41	<0.01**
Random effects					
Participants	–	–	1	654.00	<0.001***
Pictures	–	–	1	82.91	0.002**

Note: Number of observations = 914, number of stimuli = 32, number of participants = 29; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test. Objective measures did not predict average fixation counts in the 6-s condition (experiment 2), with pleasantness and complexity as the only fixed effects.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

TABLE 2 Results of linear mixed-effects model with subjective and objective fixed factors predicting average fixation counts when viewing representational paintings for 25 s

Fixed effects	*B (SE)*	95% CI	*df*	χ^2	*p(LRT)*
(Intercept)	67.03 (1.59)	[63.88, 70.17]			
Perimeter detection	1.18 (0.53)	[0.20, 2.15]	1	5.47	0.019*
Entropy	1.74 (0.66)	[0.51, 2.96]	1	7.29	0.007**
High spatial frequency	−1.91 (0.67)	[−3.15, −0.66]	1	8.38	0.004**
Complexity	2.56 (0.85)	[0.98, 4.13]	1	9.28	0.002**
Pleasantness	−2.12 (0.48)	[−3.01, −1.23]	1	17.37	<0.001***
Random effects					
Participants	–	–	1	634.08	<0.001***
Pictures	–	–	1	28.09	<0.001***

Note: Number of observations = 920, number of stimuli = 32, number of participants = 29; *B* = estimated regression coefficients (standardized predictors); CI = profile confidence intervals of regression coefficients; LRT = likelihood-ratio test. The results for the 6-s condition (experiment 2) were nearly identical, with the exception that the Canny algorithm—and not the perimeter detection measure—positively predicted fixation counts.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

a more pronounced positive effect of complexity on fixation counts ($B(SE)$ = 0.58(0.28), 95% CI = [−0.64, −0.06]) and fixation durations ($B(SE)$ = −12.16(5.45), 95% CI = [−22.79, −1.53]). Saccade amplitudes were not predicted by any of the factors under consideration (Table S16). In comparison to environmental scenes, arousal did not significantly affect viewing behavior when only representational paintings were considered in the analysis, neither in the 25-s nor in the 6-s condition.

Importantly, these findings observed for environmental scenes and representational paintings hold true when a wide range of image statistics[20,29] were considered in the respective models predicting eye movements, ruling out that the effects for affective contents and complexity were due to differences in the surface structure of the images. Tables 1 and 2 present results regarding fixation counts, and Tables S17 and S18 present results predicting saccade amplitudes. We ran these models separately for environmental scenes and paintings because correlations between subjective picture ratings, notably complexity, and objective measures (Tables S19 and S20) may depend on the picture category.[29,30]

DISCUSSION

We show in two experiments that basic perceptual processes of information intake, such as those quantified by temporal and spatial eye movement parameters during early stages of visual processing, are

different for mundane images of motivational relevance and representational paintings matched for semantic and affective contents. Our main finding regarding the role of affective and compositional factors indicates that pleasantness induced by representational paintings led to a slower and deeper processing mode than when viewing pleasant environmental scenes: more aesthetically pleasing paintings led to a smaller number of fixations and longer fixation durations. For these temporal eye movement parameters, the opposite effect was observed for environmental scenes. Additionally shorter saccade amplitudes were associated with more pleasant environmental scenes. Low-level image features could not explain these observed differences.

Our secondary finding suggests that differences between the perception of environmental scenes and representational paintings also depend on the duration of the visual experience. We did not observe a main effect of picture category on eye movements under a short presentation duration (6 s), suggesting that factors such as pleasantness and complexity are driving eye movements in both picture categories. However, when the presentation duration was much longer (25 s), paintings were explored in greater details than environmental scenes, and an additional effect of emotional arousal on eye movements emerged.

Picture complexity had a similar effect on gaze patterns across environmental scenes and representational paintings, indicating that informative objects receive more fixations.[20] In contrast, as predicted, induced pleasantness led to consistent differential effects across picture categories under both short and long presentation durations. Previous research[20] has shown that pleasant and unpleasant environmental scenes yielded more fixations in comparison to neutral pictures presented for 6 s. A recent study (not employing a free-viewing task) has reported a higher frequency of fixations for unpleasant scenes and those that were highly arousing.[26] Here, our data indicate that for a presentation duration of 6 s, as well as for a duration of 25 s, an increase in pleasantness was associated with a larger number of fixations and smaller saccade amplitudes when viewing environmental scenes, suggesting that local information from the visual array was encoded in greater detail. We do not think that this local processing mode due to high pleasantness is comparable to the ambient processing mode (i.e., short fixations and long saccades), which has been reported during visual explorations of scenes.[46]

Asymmetries in pleasantness processing when viewing environmental scenes are generally not easy to explain and are actively debated in the field.[47] A high frequency of fixations in response to unpleasant high-arousing environmental scenes is usually interpreted as a negativity bias, based on the argument that unpleasant scenes drive eye movements and information gathering because they signal threat and activate the defensive system.[26,48] However, other studies involving motivational environmental scenes[20,49] have not provided support for this interpretation, which clearly calls for further studies. Here, we observed longer fixation durations and saccade amplitudes for unpleasant environmental scenes in comparison to pleasant scenes, which is also going against the concept of negativity bias.

One difference between the present work and previous studies that varied pleasantness of environmental scenes[20,26] is that the semantic contents of the IAPS pictures differed considerably. Unlike other studies, we excluded contents, such as erotic scenes, which are often used to evoke high arousal and pleasantness, as well as highly unpleasant semantic contents, such as mutilated, burned bodies or suffering animals. This was done first to avoid exposing participants to such contents for a long presentation duration, and second to obtain similar semantic contents across the two picture categories. The latter was more achieved by avoiding extreme affective ratings, which are usually induced by the semantic contents described above. However, future studies may want to include semantic contents generating stronger affective responses, perhaps using a relatively short presentation duration so that participants are not feeling too uncomfortable when viewing such contents. For environmental scenes presented for 25 s, and also for paintings in the LMM comparing both picture categories, arousal emerged as an additional independent positive factor, which is in line with a recent eye movement study[26] with neurophysiological evidence[25] showing that pleasantness and arousal are independent affective components. However, the relative contribution of pleasantness and arousal, as well as the direction of their effects on eye movements, is still currently debated.[26] Our results suggest that arousal may emerge as a relevant factor only under longer presentation durations because pleasantness and complexity appear to be the dominant factors in the first few seconds when unfamiliar scenes are explored. Overall, our findings are consistent with the interpretation of motivated vigilance when viewing environmental scenes, in which defensive or appetitive motivation leads to goal-oriented behavior.[20,31]

In the case of visual art, however, we interpret the slower temporal information intake (i.e., fewer and longer fixations) due to increased pleasantness as indicative of a specific mental state that may well be characterized by disinterested pleasure or aesthetic distance. In contrast to environmental scenes, pleasantness had no effect on spatial parameters, such as saccade amplitudes, which further supports the view that pleasantness influences eye movement parameters differently in the two image categories. Moreover, arousal, often described as emotional intensity, also seems to play a lesser role in determining ocular responses to artworks in comparison to environmental scenes.

We also examined how presentation duration affected eye movement behavior across environmental scenes and representational paintings, hypothesizing that aesthetic experiences are more likely to occur under longer presentation durations. In the 6-s condition, fixation counts/durations and saccade amplitudes were generally not dependent on the picture type. In the 25-s condition, however, paintings were explored in greater detail than environmental scenes, as indicated by a higher number of fixations (saccade amplitudes remained unaffected). It is possible that interest, searching for meaning in art,[28] and a focus on the properties of the paintings[50] may underlie this process, which could be interpreted as a different type of motivational behavior that is not related to utilitarian goals. Our findings, based on continuous measurements of eye movements in a free-viewing task,

thus empirically support the long-standing concepts of disinterested pleasure and aesthetic distance.[5]

One limitation of our study is that we did not assess states of consciousness, such as state absorption and state flow, after participants viewed the stimuli. Future research will have to meet the challenge of determining how states of disinterested pleasure can be differentiated from altered states of consciousness, such as absorption[51] and flow,[52] which have been associated with aesthetic experiences. We initially focused on a wide range of stimulus characteristics and studied how they influenced eye movements. Based on these findings, we believe that the study of individual differences may be a fruitful avenue to characterize these mental states in greater depth in the future. Related eye movement research in a gallery setting has recently shown that eye movement parameters can mediate the relationship between personality traits and verbal responses to artworks.[53] Another method for further elucidating our findings is to employ control tasks that vary in terms of engagement in addition to a free viewing task, and further, to incorporate other factors, such as beauty and the sublime, into the research design. The latter is particularly relevant when testing other philosophical assumptions about aesthetic experiences, such as whether beauty requires thought.[54] Finally, we would encourage researchers to compare eye movements when viewing other types of aesthetic and everyday objects with a more neutral affective content, such as aesthetic landscape photographs versus common environmental scenes (e.g., neutral IAPS pictures), or when viewing design objects versus common everyday objects. Only then, we can better understand how pleasantness, arousal, and complexity determine eye movements.

Taken together, our findings suggest that measuring and analyzing temporal and spatial eye movement parameters and their relation to pleasantness, arousal, and complexity is a valuable avenue for uncovering the perceptual mode underlying aesthetic experiences. Since eye movements have been linked to brain activity,[24] these findings may have implications for the field of neuroaesthetics, in which disinterested pleasure has been hypothesized to be linked to the reward system and the differentiation between liking and wanting.[12]

ACKNOWLEDGMENTS

The authors would like to thank Nina Jahrmann, Denise Ramsperger, Lars-Michael Schöpper, and Milan Andrejevic for assistance in data collection and Bruno Gingras for valuable comments on an earlier draft of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

M.M.M. Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, visualization, and project administration. H.L. Supervision, resources, writing—reviewing and editing.

DATA AVAILABILITY STATEMENT

The data of the main analyses are available at https://osf.io/3mxk8/?view_only=6a0cb938977344368803f94cfb6588c5.

ORCID

Manuela M. Marin https://orcid.org/0000-0002-9657-1507
Helmut Leder https://orcid.org/0000-0003-3219-3671

REFERENCES

1. Nadal, M., & Chatterjee, A. (2019). Neuroaesthetics and art's diversity and universality. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*, 1487.
2. Silver, H. R. (1979). Ethnoart. *Annual Review of Anthropology*, *8*, 267–307.
3. Zaidel, D. W., Nadal, M., Flexas, A., & Munar, E. (2013). An evolutionary approach to art and aesthetic experience. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 100–109.
4. Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, *11*, 265–279.
5. Cupchik, G. C. (2002). The evolution of psychical distance as an aesthetic concept. *Culture & Psychology*, *8*, 155–187.
6. Cupchik, G. C., & Winston, A. S. (1996). Confluence and divergence in empirical aesthetics, philosophy, and mainstream psychology. In M. P. Friedman & E. C. Carterette (Eds.), *Handbook of perception & cognition, cognitive ecology* (pp. 62–85). San Diego, CA: Academic Press.
7. Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*, *33*, 239–251.
8. Gerger, G., Leder, H., & Kremer, A. (2014). Context effects on emotional and aesthetic evaluations of artworks and IAPS pictures. *Acta Psychologica*, *151*, 174–183.
9. Dissanayake, E. (2000). *Art and intimacy. How the arts began*. Seattle and London: University of Washington Press.
10. Hayn-Leichsenring, G. U., & Chatterjee, A. (2019). Colliding terminological systems—Immanuel Kant and contemporary empirical aesthetics. *Empirical Studies of the Arts*, *37*, 197–219.
11. Dickie, G. (1964). The myth of the aesthetic attitude. *American Philosophical Quarterly*, *1*, 56–65.
12. Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, *1369*, 172–194.
13. Marin, M. M. (2015). Crossing boundaries: Toward a general model of neuroaesthetics. *Frontiers in Human Neuroscience*, *9*, 443.
14. Leder, H., Gerger, G., Brieber, D., & Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cognition and Emotion*, *28*, 1137–1147.
15. Van Dongen, N. N. N., Van Strien, J. W., & Dijkstra, K. (2016). Implicit emotion regulation in the context of viewing artworks: ERP evidence in response to pleasant and unpleasant pictures. *Brain and Cognition*, *107*, 48–54.
16. Moran, T. P., Jendrusina, A. A., & Moser, J. S. (2013). The psychometric properties of the late positive potential during emotion processing and regulation. *Brain Research*, *1516*, 66–75.
17. Cupchik, G. C., Vartanian, O., Crawley, A., & Mikulis, D. J. (2009). Viewing artworks: Contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain and Cognition*, *70*, 84–91.
18. Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2013). When we like what we know – A parametric fMRI analysis of beauty and familiarity. *Brain and Language*, *124*, 1–8.
19. Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response to facial attractiveness. *Neuropsychology*, *23*, 135–143.

20. Bradley, M. M., Houbova, P., Miccoli, L., Costa, V. D., & Lang, P. J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, *48*, 1543–1552.
21. Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*, 6.
22. Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, *11*, 17.
23. Cronin, D. A., Hall, E. H., Goold, J. E., Hayes, T. R., & Henderson, J. M. (2020). Eye movements in real-world scene photographs: General characteristics and effects of viewing task. *Frontiers in Psychology*, *10*, 2915.
24. Simola, J., Fevre, K. L., Torniainen, J., & Baccino, T. (2015). Affective processing in natural scene viewing: Valence and arousal interactions in eye-fixation-related potentials. *Neuroimage*, *106*, 21–33.
25. Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*, 247–265.
26. Gomez, P., von Gunten, A., & Danuser, B. (2019). Eye gaze behavior during affective picture viewing: Effects of motivational significance, gender, age, and repeated exposure. *Biological Psychology*, *146*, 107713.
27. Unema, P. J., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*, 473–494.
28. Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 489–508.
29. Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS One*, *8*, e72412.
30. Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica*, *163*, 38–58.
31. Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical report A-6. University of Florida.
32. Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience*, *10*, 536.
33. Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.
34. Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. J. G. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) [Multidimensional mood questionnaire]*. Göttingen: Hogrefe.
35. Specker, E., Forster, M., Brinkmann, H., Boddy, J., Pelowski, M., Rosenberg, R., & Leder, H. (2020). The Vienna Art Interest and Art Knowledge Questionnaire (VAIAK): A unified and validated measure of art interest and art knowledge. *Psychology of Aesthetics, Creativity, and the Arts*, *14*, 172–185.
36. Smith, N. D., Glen, F. C., & Crabb, D. P. (2012). Eye movements during visual search in patients with glaucoma. *BMC Ophthalmology*, *12*, 45.
37. Delplanque, S., N'diaye, K., Scherer, K., & Grandjean, D. (2007). Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods*, *165*, 144–150.
38. Bates, D., Mächler, M., Bolker, B., & Walker, S. J. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
39. Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
40. Judd, C. M., Westfall, J., & Kenny, D. A. (2021). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54–69.
41. Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*. Washington, DC: Sage Publications.
42. Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.
43. Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.
44. Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
45. Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138.
46. Pannasch, S., Helmert, J. R., Roth, K., Herbold, A. K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*. https://doi.org/10.16910/jemr.2.2.4
47. Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*, 383–403.
48. Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.
49. Dietz, J., Bradley, M. M., Okun, M. S., & Bowers, D. (2011). Emotion and ocular responses in Parkinson's disease. *Neuropsychologia*, *49*, 3247–3253.
50. Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
51. Vessel, E. A., Starr, G. G., & Rubin, N. (2012). The brain on art: Intense aesthetic experience activates the default mode network. *Frontiers in Human Neuroscience*, *6*, 66.
52. Csikszentmihalyi, M., & Robinson, R. E. (1990). *The art of seeing: An interpretation of the aesthetic encounter*. Report No. ISBN-0-89236-156-5. Los Angeles, CA: Getty Center for Education in the Arts.
53. Trawiński, T., Palumbo, L., Harrison, N. R., Kass, J., Metelmann, A. C., Bari, S., & Donnelly, N. (2021). Aesthetic attention: What is the role of eye movements, personality and cognitive style on responses to artworks? *Perception*, *50*(1_suppl), 56.
54. Brielmann, A. A., & Pelli, D. G. (2017). Beauty requires thought. *Current Biology*, *27*, 1506–1513.e3.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Marin, M. M., & Leder, H. (2022). Gaze patterns reveal aesthetic distance while viewing art. *Ann NY Acad Sci.*, *1514*, 155–165. https://doi.org/10.1111/nyas.14792

2.6.4 Summary and outlook

We employed linear mixed-effects modelling to predict fixation durations, fixation counts and saccade amplitudes by complexity, arousal, pleasantness and picture category. In Experiment 1, the presentation duration was 25 s and participants viewed both environmental scenes and representational paintings. We observed a significant effect of picture category, suggesting that paintings received on average a higher number of fixations compared to environmental scenes. There was also an effect of arousal, and high-arousing pictures received more fixations than low-arousing ones. Saccade amplitudes were longer for pleasant than for unpleasant paintings, whereas the reverse held true for environmental scenes. Interestingly, complexity was not a significant predictor in any of the three models predicting eye movement parameters.

For Experiment 2, based on a shorter presentation duration of 6 s, a similar analysis revealed that picture category was not a significant factor, but instead, visual complexity was: more complex images received more fixations. Pleasantness interacted with picture category, and scanning behaviour decreased (i.e., fewer fixations and longer fixation durations) with an increase in pleasantness when viewing paintings. There was no effect of arousal, and the model for saccade amplitudes did not reveal any significant fixed effects.

Due to several (nearly) significant interactions between pleasantness and picture category we also analyzed the data for each picture category separately. Our clear pattern of results suggests that the factor of pleasantness, and not visual complexity, leads to opposite effects on eye movement parameters when comparing environmental scenes with representational paintings: fixation durations were longer for pleasant paintings and shorter for pleasant environmental scenes. We also found that saccade amplitudes were shorter for pleasant environmental scenes; however, saccade amplitudes were not affected by complexity and pleasantness when viewing representational paintings. Arousal was only a predictor of eye movement parameters when viewing environmental scenes for 25 s. Importantly, these findings held true when image features related to complexity and affect were considered in the statistical models. Altogether, we believe that our findings are solid because they were obtained in a within-subjects design (picture type) and replicated with images presented for 6 s in a different sample.

While early studies suggested that eye movements when viewing environmental scenes in a free viewing task are dependent on complexity and pleasantness (Bradley et al., 2011), our results suggests that arousal is also a significant factor for a long presentation duration of 25 s (see also Löw et al., 2013; Simola et al., 2015). We would encourage researchers to examine several presentation durations between 6 and 25s to see when arousal becomes a contributing factor besides complexity and pleasantness in environmental scene perception. Arousal is related to alertness and excitement and as such often an indicator of action readiness which enhances survival (Reisenzein, 1994). Future studies could measure psychophysiological responses as well as event-related potentials (Cuthbert et al., 2000; Olofsson et al., 2008) to get a better understanding of the role of arousal in environmental scene perception, especially in comparison to the processing of representational paintings for which we did not observe any arousal effects on eye movements. It is possible that although variations of emotional arousal are present at the subjective and psychophysiological level when viewing visual art, eye movements may remain unaffected. This hypothesis could easily be tested in the future by combining, for example, eye movement measurements with the study of event-related potentials (Simola et al., 2015).

In the context of Berlyne's (1971) psychobiological model, our findings demonstrate that complexity and pleasantness seem to be variables that influence ocular behaviour more than arousal. However, from this one cannot conclude that arousal is an unimportant factor when viewing environmental scenes and representational paintings in general. It is possible that arousal may play a minor role at the behavioural level as assessed by eye movements, but that it is still important at the subjective level. Future studies will have to look into what types of behavioural measures are affected by arousal and which not when viewing affective scenes of various types.

This study demonstrated, as Marin and Leder (2016) did, that it is worth considering different presentation durations when comparing environmental scenes and representational paintings to elucidate the nature of aesthetic experiences. Although differential effects of pleasantness on eye movement parameters were present under both short and long presentation durations, which we interpret as being characteristic of a specific viewing mode related to art, picture type as a significant main effect only appeared in the 25-s condition: paintings received more fixations and the more

pleasant a painting was, the longer were the saccade amplitudes. Therefore, studies involving several presentation durations may be able to track how art-specific viewing behaviour emerges over time. Subjective reports of interest could complement such studies to see whether this variable could explain the differential results for environmental scenes and artworks. Another possible explanation for a higher number of fixations when viewing artworks compared to environmental scenes is perceived cognitive challenge (Ganczarek et al., 2020). Recent advances in eye movement research have promoted the idea of categorizing the duration of eye movements according to their length and to predict aesthetic evaluations (Jankowski et al., 2020). The current data could be re-analyzed within this framework.

Eye movements seem to depend not only on the affective contents but also on social information when viewing natural scenes (Birmingham et al., 2009; Kaisler et al., 2020; Rösler et al., 2017; Skripkauskaite et al., 2023). Research using representational paintings has shown that eye movements are sensitive to social interactions depicted in the paintings (Villani et al., 2015) and that trait empathy of the viewer is also a relevant factor (see also Wilkinson et al., 2024). In the current study, we unfortunately did not collect any data on affect-related personality traits (see Marin & Leder, 2018) and future studies will need to consider individual differences in greater depth.

We worked with a rather limited set of 32 images for each of our two picture categories because the focus was on stimulus complexity and emotion as well as on the comparison between environmental scenes and representational paintings. However, future research will have to meet the challenge to select stimuli that vary along complexity, pleasantness and arousal as well as on some kind of social dimension (social vs. non-social). This will make the research design more complex but also more complete. Our stimuli were pre-selected on the basis of ratings obtained by two groups of people who were not participating in the eye movement study. This approach has the advantage that one can show many pictures in a free-viewing task (even with long presentation durations) without receiving subjective ratings in an additional, time-consuming block after measuring eye movements. The drawback is that individual differences (regarding affective responses, familiarity and the perception of social information in a scene) cannot be included in the analysis. It may also be fruitful to randomize environmental scenes and paintings within one block to replicate and extend the current findings. Future studies will have to find a compromise be-

tween the number of stimuli and the integration of measures of individual differences to move forward in the field.

Our findings can be discussed within the context of motivational theories of affect (Bradley et al., 2011; Niu et al., 2012) and suggest that a particular viewing mode, especially based on pleasantness induced by the image, may be able to differentiate between the perception of paintings and everyday motivational scenes. Pleasure and pleasantness have been frequently discussed within motivational theories of reward (McCrae, 2024), and it is possible that this variable is crucial for differentiating between an aesthetic viewing mode and a viewing mode serving utilitarian goals.

Since eye movements have been described as a window to the brain and cognition (König et al., 2016), we suggest that the field of neuroaesthetics would profit from integrating eye movement measurements into their research designs. For instance, Vessel et al. (2012, p. 103) reported activation of the default-mode network during the perception of highly moving paintings, and the authors interpreted this finding as “internally directed mentation” (for a review on the neural correlates of aesthetic experiences see Vessel et al., 2022). We did not collect measures of altered states of consciousness, such as flow or absorption (Husselman et al., 2024; Kuijpers et al., 2017), which may complement future neuroimaging studies paired with eye movement measurements. Especially time-sensitive methods such as electroencephalography seem to be suited for such a purpose.

Papenmeier et al. (2024) recently investigated how bottom-up (image type) and top-down factors (cognitive framing) influence gaze behaviour and aesthetic ratings of real artworks and scientific depictions. They reported that the image type determined eye movements and not framing. Actual artworks were associated with “fewer and longer fixations, a shorter scanpath, a shorter distance of fixations to the screen centre, and a lower similarity of viewing patterns across participants” (p. 642). They did not relate their subjective ratings to the eye movement analysis, so it is not possible to directly compare their results with ours, but Papenmeier et al. (2024) also reported a slower information processing mode when participants were viewing real artworks, attributing these effects to the visual appearance and not to a specific cognitive art schema. Future studies should aim to replicate these and our findings with other types of visual artworks.

2.7 The role of collative variables in aesthetic experiences (Marin, 2022)

2.7.1 Introduction and summary

This book chapter (Marin, 2022) sought to present a comprehensive overview of Berlyne's theory of collative motivation (1971) and its influence on the field of empirical aesthetics up to the present. Specifically, the chapter focuses on research involving music, the visual arts and literature/poetry. Following a detailed description of Berlyne's theory in its historical context, the current relevance of this theory to the field of empirical aesthetics is considered. Among other topics, the frequent reduction of Berlyne's theory to the inverted-U relationship between collative variables and hedonic tone, the apparent neglect of arousal when studying this relationship, as well as the comprehensive research approach proposed by the *New Experimental Aesthetics* are discussed. The next section presents Berlyne's theory within the context of reward, pleasure, motivational arousal and the brain. The progress in the fields of cognitive and affective neuroscience of the last 25 years has markedly changed our view of the neural basis of Berlyne's theory (i.e., the arousal systems he proposed). Then, four sections are dedicated to research on each of the four main collative variables proposed by Berlyne, namely *novelty-familiarity*, *surprise-expectedness*, *complexity-simplicity* and *ambiguity-clarity*. By including behavioural and neuroscientific studies involving music, the visual arts and literature/poetry, these sections offer a broad view on the latest developments regarding Berlyne's theory. For instance, it becomes apparent that certain collative variables are more frequently studied with regard to music than with regard to other art forms. The chapter ends with reflecting on major challenges, goals and suggestions for future research.

This literature review chapter was written towards the end of the current doctoral thesis project, thus it includes articles of this thesis that were already published. It offers an overview and introduction to current research related to Berlyne's theory and demonstrates how articles related to this thesis are well-embedded in the wider literature on this topic.

Published as part of *The Oxford Handbook of Empirical Aesthetics* (Nadal & Vartanian, 2022), this chapter will hopefully become a valuable resource for students and researchers being interested in the legacy of Berlyne's research output. Thanks to its wide scope, this chapter will allow researchers to gain a comprehensive understanding of how Berlyne's theory has influenced the research field in different ways over the last decades, and which aspects of Berlyne's theory are still considered as

valid today. Therefore, it will also help researchers discuss the evidence in favour of Berlyne's theory in relation to other existing theories on aesthetic experiences. Finally, the explicit comparisons between respective art forms may inspire researchers to pursue a comparative approach to empirical aesthetics in the long run, which may accelerate the progress of the field.

2.7.2 Authorship contribution statement

MMM wrote the book chapter.

2.7.3 Re-print of the published book chapter

Chapter 18

The Role of Collative Variables in Aesthetic Experiences

Manuela M. Marin

Introduction

The systematic study of stimulus properties and their role in aesthetic responses can be traced back to the birth of empirical aesthetics, namely to Gustav Theodor Fechner's (1876) ground-breaking publication *Vorschule der Ästhetik*. Fechner describes a set of aesthetic principles of which some largely depend on object-related factors (e.g., the aesthetic threshold, which reflects his psychophysical approach), whereas other principles depend more on person-related factors (e.g., associations). It is not surprising that Fechner acknowledged the mutual role of stimulus properties and the perceiver in these early days of empirical aesthetics because they constitute essential pillars of aesthetic communication already known in antiquity: Greek architects were aware that corrections needed to be made to the canonical order of temples (taking into account visual perception) to obtain a regular, balanced, and pleasing result in the perceiver.

In the 20th century, psychological theories on aesthetic experiences and behavior ascribed a differential role to object- and person-related factors, generally depending on the psychological tradition in which they were developed. For example, stimulus properties were central to the work of Birkhoff (1932) and Eysenck (1941), who tried to predict an aesthetic response using mathematical formulae that were based on visual order and complexity. Stimulus properties were also at the core of subsequent theories that were influenced by information theory (Shannon, 1948). A case in point is Daniel Berlyne's (1924–1976) seminal arousal-based theory of aesthetic experience and behavior (Berlyne, 1960, 1971, 1974c). Berlyne, a motivational psychologist, greatly advanced the field of empirical aesthetics from around 1960 until his death in 1976 by introducing new approaches, methods, and aims, which led to a movement commonly referred to as the

New Experimental Aesthetics (Berlyne, 1971, 1974a). Berlyne and his numerous followers concentrated on (1) *collative* (structural, formal) stimulus properties, (2) motivational questions, (3) the measurement of nonverbal behavior and subjective judgments, and finally, (4) they tried to contribute to a better understanding of human psychology through the study of aesthetic phenomena (Berlyne, 1974a). Cupchik (1988) considered Berlyne as the "founder of modern experimental aesthetics" (p. 171), pointing toward Berlyne's reductionist and experimental approach, which gave a strong impetus to the research community in the 1970s and thereafter (Day, 1981). Consequently, Berlyne's *psychobiological model* (sometimes simply referred to as *arousal theory* or the *two-factor theory*) is widely known and frequently cited in reviews on the history and development of empirical aesthetics up to the 21st century (Palmer, Schloss, & Sammartino, 2013; Silvia, 2012; Vartanian, 2014).

Berlyne's research was largely guided by the question of why humans "display curiosity and explore their environment, why they seek knowledge and information, why they look at paintings and listen to music, [and] what directs their train of thought" (Konečni, 1978, p. 134). His answers to some of these fundamental questions were described in his book *Conflict, Arousal and Curiosity* (1960), which culminated in his general theory of *collative motivation*. This theory was later applied to the arts in his influential monograph *Aesthetics and Psychobiology* (1971). The intellectual achievement of Berlyne's work can be attributed to the introduction of an affective component (arousal) to the study of perceived stimulus features and their role in aesthetic responses. Specifically, he brought together aspects of Clark Hull's (1943) *drive reduction theory* (see also Bolles, 1975, for the history of its demise) with aspects of *information theory* by integrating them with the latest neurophysiological findings on reward processing of that time (Olds & Olds, 1965). Clearly, Berlyne's theoretical work was largely rooted outside aesthetics and Walker states that "it was Berlyne's ambition to achieve a truly general and universal theory, applicable to many, if not all, aspects of human activity" (1980, p. 110). Berlyne (1971) presented his theoretical arguments in a rigorous way and captured readers' attention by offering convincing examples stemming from music, poetry, and the visual arts alike. Berlyne was thus able to approach the study of empirical aesthetics from a broad perspective (Machotka, 1980)—a scientific approach that is nowadays not frequently followed (but see Brown, 2019; Marin, 2015; Tiihonen, Brattico, Maksimainen, Wikgren, & Saarikallio, 2017). It is worth noting that Berlyne's academic contributions and active scholarly involvement were already highly appreciated during his lifetime (Cupchik, 1988; Konečni, 1978; Machotka, 1980).

In brief and simplified terms, Berlyne (1960, 1971) argued that a stimulus possesses *arousal potential*, which he later described as the "overall power to excite the nervous system" (Berlyne & Madsen, 1973, p. 14), suggesting a very broad perspective on physiological arousal. Arousal potential is determined additively by three types of variables: *psychophysical* (e.g., intensity, brightness, and color), *ecological* (e.g., signal value, association with biologically significant events), and *collative* (e.g., novelty, complexity, and ambiguity). The term "collative variables" was coined by Berlyne because these variables "involve comparison, and thus response to degree and nature of similarity or difference, between stimulus elements that may be present together or at different times" (1971, p.

141). According to Berlyne (1960), collative variables are the primary drivers of arousal changes, which in turn are associated with changes in *hedonic value* (e.g., pleasure, preference, and beauty).

Berlyne's *arousal-boost* mechanism has been most frequently studied in the field and underlies the well-known inverted U-shaped curve, which is a function of a collative variable and a measure of hedonic value (Figure 18.1). It assumes that the perceiver is in a moderate arousal state *prior* to the perception of a stimulus; however, this person-related state is often ignored (but see, e.g., Konečni, 2011, for considering the immediate arousal context in relation to music choice). According to Berlyne (1971), activation of the primary reward system by a stimulus with moderate arousal potential may lead to maximum hedonic value. When the arousal potential of a stimulus is too large, hedonic value decreases and becomes negative due to activation of the aversion system, which has a higher activation threshold than the primary reward system. The *arousal-jag* mechanism, less known but still relevant, describes how positive hedonic value can be achieved in a case when arousal first increases markedly (high state of tension) but is then reduced, leading to relief from unpleasantness. In this case, the secondary reward system is active and there is an inhibition of the aversion system. This short description illustrates the core of Berlyne's theory by suggesting that the perception of object-related properties influences the arousal level of a person, which in turn motivates behavior (exploration, approach, and avoidance) and subjective affective experience (e.g., pleasure).

Berlyne (1965) argued that a work of art consists of a group of elements that can communicate four types of information, namely semantic, expressive, cultural, and syntactic, of which the latter three types can be grouped into *aesthetic information*. In abstract art forms, such as music and abstract paintings, syntactic information is the primary source of information (but see Koelsch, 2009; Koelsch et al., 2004, for

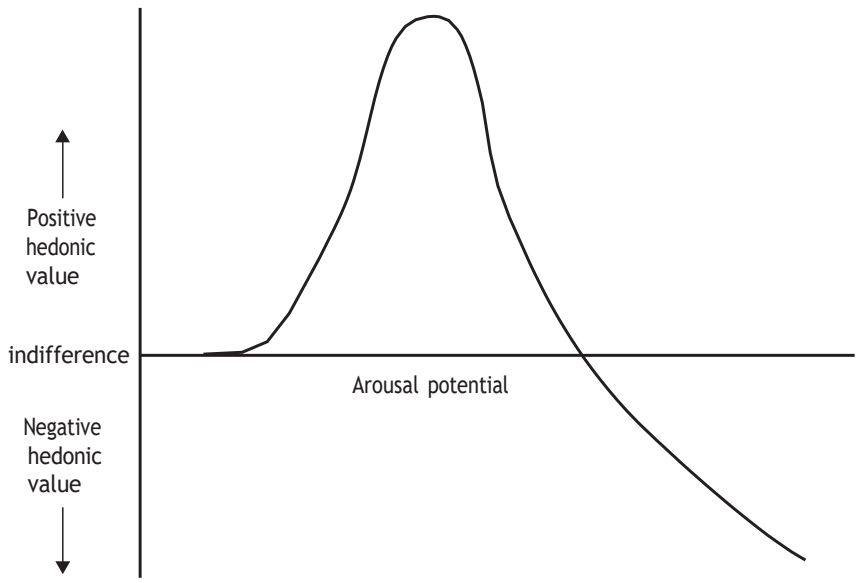

Figure 18.1. Resulting hedonic tone as a function of arousal potential, as discussed in Berlyne, 1971. (First developed by Wilhelm Wundt.)

neuroscientific evidence of how music can also convey semantic meaning). But even in works of art with semantic content, collative stimulus properties are crucial because artistic communication, within Berlyne's framework (1974), is largely determined by *how* something is communicated and not by *what* is communicated. In the *synthetic* research approach, Berlyne and collaborators constructed stimulus patterns that varied systematically along collative variables, whereas in the *analytic* approach, they studied actual artworks. These two empirical approaches, dating back to Fechner (Berlyne, 1971), are still pursued in parallel today.

Studying collative variables and their relation to aesthetic experience and behavior has received wide attention in the field since the 1970s—in and beyond the context of Berlyne's theory of collative motivation. Unsurprisingly, then, variables such as novelty, complexity, ambiguity, and so on also play a direct or indirect role in subsequent models of aesthetic experience (Carbon & Jakesch, 2013; Graf & Landwehr, 2015; Leder, Belke, Oeberst, & Augustin, 2004; Pelowski, Markey, Forster, Gerger, & Leder, 2017b; Van de Cruys & Wagemans, 2011). Arousal, regardless of its underlying (frequently vaguely defined) concept, is in most cases not regarded as an important factor in these models. Beyond the field of basic psychological research, collative variables have been of interest to those studying website aesthetics (Post, Nguyen, & Hekkert, 2017), online marketing (Sohn, Seegebarth, & Moritz, 2017; Wang & Li, 2017) brand logo design (Wang Duff, & Clayton, 2018), fashion (Tseng, 2018), and interior architectural design (Yeun, Eunsoo, & Jung, 2018), to name just a few examples from the visual domain. Collectively, this demonstrates the theoretical significance of collative variables and their practical relevance in aesthetic communication.

This chapter will critically review research on the role of collative variables in aesthetic experiences and behavior—within Berlyne's theoretical framework and beyond—as well as present challenges faced by the fields of psychology and cognitive neuroscience. Beginning with a more detailed presentation of Berlyne's theory from the perspective of current research in empirical aesthetics, insights gained from neuroaesthetics related to reward, motivational arousal, and pleasure will be presented. Next, findings and issues regarding individual collative variables (novelty, surprise, complexity, and ambiguity) will be discussed in more detail. Whenever appropriate, the examples will draw on various subfields of empirical aesthetics, mostly originating from the study of different types of visual stimuli, music, and poetry, which ultimately allows for a broad view on collative variables across the arts. Finally, this chapter concludes with a discussion of major challenges, goals, and suggestions for applications of current research findings regarding the role of collative variables in aesthetic experiences.

Current Perspectives on Berlyne's Theory of Collative Motivation

Berlyne's psychobiological model (1971) is often reduced to his proposed inverted U-shaped curve, which does not do full justice to his contribution to the field. Berlyne not

only tried to explain *how* stimulus properties are linked to hedonic responses, but he also provided an explanation for *why* humans engage with works of art although this behavior has no immediate survival value. Information-seeking, learning, and motivation are at the core of Berlyne's answers to this vital question when regarding animals and humans as curious beings that, depending on the context and their inner state, can show specific types of *curiosity*. Berlyne (1954) defined *perceptual curiosity* as being evoked by a complex or ambiguous sensory stimulus pattern that motivates exploration to gain information. *Epistemic curiosity* is evoked by complex ideas or conceptual ambiguities (e.g., scientific theories) that lead to further cognitive efforts to acquire knowledge, such as hypothesis testing. In this context, Berlyne further differentiates between *specific* and *diversive exploratory behaviors*. Specific exploration (e.g., due to uncertainty and conflict evoked by a stimulus and driven by both types of curiosity) may be rewarding through arousal reduction, whereas diversive exploration (due to boredom) may be rewarding through a moderate increase in arousal when searching for new information. Most current theoretical frameworks do not consider curiosity as a central concept driving aesthetic behavior (Pelowski, Markey, Lauring, & Leder, 2016); however, art museum visitors are clearly motivated by curiosity and learning among other things (for a review, see Pelowski, Forster, Tinio, Scholl, & Leder, 2017a; Slater, 2007). Thus, future theoretical work, based on the latest neuroscientific findings, may need to reconsider the role of motivation in aesthetic responses (see Silvia, 2005a).

The concept of *hedonic value*, as described and used by Berlyne with the background of learning, instrumental conditioning, and reward in mind, deserves some attention since it embraces a set of variables that can be measured either through verbal judgments or nonverbal behavior. *Positive hedonic value* refers to "pleasure, reward value, positive feedback, attractiveness (i.e., capacity to elicit approach), and positive incentive value," whereas *negative hedonic value* comprises "unpleasantness, punishment value, negative feedback, repulsiveness, and negative incentive value" (Berlyne, 1971, pp. 80–81). Berlyne (1974a) further introduced the term *hedonic tone* to refer to verbal judgments of pleasure, preference, beauty, and the like, and he was particularly interested in correlations between hedonic tone and nonverbal measures of hedonic value (e.g., Berlyne, 1972; Berlyne & Lawrence, 1964). Marin, Lampatz, Wandl, and Leder (2016) demonstrated that although measures of hedonic tone (beauty, pleasantness, and liking) are strongly correlated with each other when either viewing representational paintings or listening to music, their respective relationship with a collative variable, such as complexity, may differ. This suggests that the concept of hedonic tone may be less uniform than previously assumed, which should be taken into account in future experimental work (for a differentiation between concepts of hedonic value in current models of aesthetic experience, see, e.g., Brattico, Bogert, & Jacobsen, 2013; Pelowski et al., 2017b).

Berlyne's comprehensive research approach is nowadays not common, which is not only due to the emergence of neuroscientific methods. Researchers in Berlyne's tradition initially assessed a wide range of verbal ratings (descriptive, evaluative, internal-state, and stylistic scales) using the semantic differential, measured psychophysiological arousal, and employed behavioral measures of exploration time and exploratory choice (Berlyne, 1974c). They also measured interestingness (or interest, see also Silvia, 2008,

for defining *interest* as a knowledge emotion in the context of appraisal theory) besides pleasantness (or *pleasingness*, a term previously used) as an evaluative judgment. Interest was understood as a measure of behavioral exploration, whereas pleasingness was related to reward (Saklofske, 1975). Nowadays, researchers usually adhere to only a few measures within the context of a study, such as one collative variable and one measure of hedonic tone (e.g., Bohrn, Altmann, Lubrich, Menninghaus, & Jacobs, 2013; Cui, Collett, Troje, & Cuddy, 2015; Forsythe, Nadal, Sheehy, Cela-Conde, & Sawey, 2011). Studies considering two collative variables simultaneously are rare (Heyduk, 1975; Jakesch & Leder, 2015; Madison & Schiolde, 2017; Seifert & Chattaraman, 2017; Tinio & Leder, 2009), but interesting interaction effects on hedonic value may exist as already reported by Berlyne (1970) regarding visual complexity and novelty. Moreover, partly due to this reductionist approach when studying aesthetic responses, there is still an ongoing discussion about the relative role of collative variables (in comparison with other types of factors) in responses to music (North & Hargreaves, 2000; Schäfer & Sedlmeier, 2010; Schubert, 2010), visual stimuli (Hekkert & Leder, 2008; Martindale, Moore, & Anderson, 2005; Moshagen & Thielsch, 2010; Stich, Eisermann, Knäuper, & Leder, 2007; Tuch, Presslaber, Stöcklin, Opwis, & Bargas-Avila, 2012), and literary devices (Jacobs & Kinder, 2018). To clarify this debate will be a challenging endeavor requiring rigorous experiments in the future, which will need to acknowledge the context in which aesthetic experiences happen to a much greater extent, as shown by recent studies conducted in the museum (Brieber, Nadal, & Leder, 2015; Brieber, Nadal, Leder, & Rosenberg, 2014).

There is a conceptual lack of clarity surrounding arousal, and its measurement has been increasingly neglected in the field of empirical aesthetics. This may be either due to the fact that Berlyne's original notion of arousal has considerably changed since his theory was developed, or due to a general loss of interest in studying (motivational and emotional) arousal. For instance, even within the framework of Berlyne's theory the measurement of arousal changes is currently being ignored (e.g., Chmiel & Schubert, 2018; North & Hargreaves, 1995; Sun, Yamasaki, & Aizawa, 2018). In a similar fashion, reviews on the inverted U-shaped relationship between collative variables and hedonic tone do not address empirical evidence for or against the mediation mechanism proposed by Berlyne in much depth (Chmiel & Schubert, 2017; Nadal, Munar, Marty, & Cela-Conde, 2010; for recent studies assessing arousal within the context of Berlyne's theory and beyond, see, e.g., Blijlevens, Carbon, Mugge, & Schoormans, 2012; Leder, Ring, & Dressler, 2013; Marin & Leder, 2013; Schäfer & Sedlmeier, 2011; Steinbeis, Koelsch, & Sloboda, 2006; Tschacher et al., 2012; Tuch, Bargas-Avila, Opwis, & Wilhelm, 2009). At the same time, arousal changes are still being studied when exploring emotional peak experiences during aesthetic processing, such as the chill response (Benedek & Kaernbach, 2011; Colver & El-Alayli, 2016), emotional lacrimation (Wassiliwizky, Jacobsen, Heinrich, Schneiderbauer, & Menninghaus, 2017a), or the feeling of being moved (Menninghaus et al., 2015). Yet in these cases arousal is not necessarily considered within a motivational framework.

Emotions are central to our understanding of aesthetic responses, and there are several aspects of Berlyne's theory (1971) that are debated today with regard to emotion

processing. When discussing arousal dynamics and reward it is important to note that Berlyne regarded motivational arousal and pleasure (pleasantness) as dependent variables. Pleasure is also closely related to the affective concept of valence (pleasantness), and Berlyne was aware of emotion models that claimed the independence of arousal and pleasantness (Bradley & Lang, 1994; Osgood, Suci, & Tannenbaum, 1957; Wundt, 1896). Interestingly, the controversy surrounding the nature of the relationship between arousal and pleasantness/pleasure is still ongoing in emotion science (e.g., Bestelmeyer, Kotz, & Belin, 2017; Kragel & LaBar, 2016; Kuppens, Tuerlinckx, Russell, & Barrett, 2013) as well as among those studying aesthetic responses and reward (Blood & Zatorre, 2001; Marin & Leder, 2013; Salimpoor, Benovoy, Longo, Cooperstock, & Zatorre, 2009; Schäfer & Sedlmeier, 2010). Furthermore, Berlyne's focus on changes in hedonic value and pleasure has been criticized for ignoring the myriad of emotional responses an aesthetic experience may entail (Konečni, 1996; Silvia, 2005b). Another related issue already mentioned above concerns the relative contribution of psychophysical, ecological, and collative variables to the experience of pleasure. For instance, many types of visual artworks communicate (strong) emotions through semantic content and color, which alongside collative variables can generate pleasure, but still little is known about how aspects of emotional content interact with collative variables during visual aesthetic experiences (Marin & Leder, 2013). Indeed, the current neurocognitive literature on emotion–cognition interaction suggests that factors such as emotional content and complexity may interact at early stages of affective picture processing (e.g., Carretié, Hinojosa, López-Martín, & Tapia, 2007; Low, Bradley, & Lang, 2013; Schlochtermeier et al., 2013; but see also Bradley, Hamby, Löw, & Lang, 2007). Altogether, such findings may provide fruitful pathways for theoretical advances in the field.

Berlyne's arousal theory has been criticized despite its large influence (Konečni, 1996), and alternative theories have been proposed, such as Martindale's prototype theory of aesthetic preference (Martindale, 1984; Martindale & Moore, 1988), Silvia's appraisal account of aesthetic emotions and interest (Silvia, 2010, 2012), and the processing fluency theory of aesthetic pleasure (Reber, Schwarz, & Winkielman, 2004). Berlyne's theory has neither been further developed nor been integrated into other theoretical accounts, which is surprising because researchers are still frequently studying the inverted U-shaped curve using behavioral measures. In more recent times, the work on the role of neurotransmitters during aesthetic experiences has replaced work on arousal.

Reward, Motivational Arousal, and the Brain

In his theoretical work, Berlyne (1971) focused much on arousal modulation in areas of the *ascending reticular activating system* (ARAS) in the brainstem (Moruzzi & Magoun, 1949). He assumed, mostly based on animal research, that the brain areas controlling fluctuations in arousal were the same as the ones processing reward and pleasure, or that

these areas were at least close to each other. As discussed above, his theory on arousal fluctuation builds on the idea that there are three subcortical reward systems interacting with each other, two in the reticular activating system and one at the upper brainstem and in the limbic system. The notion of an involvement of several brain regions during aesthetic experiences is still valid today; however, Berlyne's conception of the reward and arousal systems is outdated.

Brain imaging has considerably changed our understanding of the human reward and arousal systems. Nowadays it is known that there is no uniform arousal system in humans and animals that can explain wake–sleep states, consciousness, as well as heightened attentional and general affective processing (Andretic, van Swinderen, & Greenspan, 2005; Lee & Dan, 2012; Picard, Fedor, & Ayzenberg, 2016; Robbins et al., 2006). The complexity of the neural underpinnings of general emotion processing in humans, for example, can be demonstrated by discussing the role of the brainstem (Venkatraman, Edlow, & Immordino-Yang, 2017), which is the oldest brain structure present in all vertebrates. The brainstem and its nuclei are particularly relevant in the discussion of appetitive motivation and motivational arousal associated with dopamine release (Wise, 2004).

The modulatory neurotransmitter network (another term for ARAS), projects to the hypothalamus, cortex, and forebrain and it controls arousal and consciousness by modulating interactions between the ascending and descending networks. This modulatory network comprises three major areas of dopaminergic neurons in the midbrain, with three different ascending pathways, that show both tonic and phasic firing patterns related to reward and motivational arousal. In addition, there are serotonergic neurons in the raphe nuclei that are functionally related to anxiety, stress, depression, and social behavior. Norepinephrine signaling in the locus coeruleus is relevant for the processing of emotionally salient information and alertness, and, together with dopaminergic pathways, it plays a role in learning and emotional memories. Moreover, cholinergic cell groups are part of the modulatory network and seem to play a role in emotion and reward processing by modulating dopaminergic signaling. It is generally agreed that the brainstem also receives input (top-down feedback) from rostral areas during bottom-up (sensory-based) emotion processing and concurrent fixed-pattern behavior (Venkatraman et al., 2017).

The above description of the brainstem circuits clearly suggests that any arousal-based theory of information processing in relation to art would need to consider the intricacy of the brainstem circuits in much more detail than Berlyne was able to do at his time. The involvement of the brainstem in a specific mechanism of emotion induction by music, the so-called *brainstem reflex*, has already been described by Juslin and Västfjäll (2008). Recent interdisciplinary models of emotion, such as the *Quartet Theory* (Koelsch et al., 2015), which is also applicable to music processing, also acknowledge the brainstem system as one of four core systems, alongside the diencephalon, hippocampal formation, and the orbitofrontal cortex. Another current model on visual art perception (VIMAP) also acknowledges the various brainstem functions at early processing stages (Pelowski et al., 2017b), but without considering the possibility of feedback loops

from higher cortical areas. Modern brain-based models go beyond the idea that (de) activations of only a few brain regions can explain complex phenomena such as aesthetic experiences, decision making, and subsequent behaviors.

Berlyne developed his psychobiological model as a broad theory, and current neuroimaging results seem to support his idea that pleasure derived from art relies on similar brain networks as other fundamental pleasures. In fact, there is abundant evidence that light, food, sex, and social contact, as well as higher pleasures, such as money, music, and visual art, activate an overlapping and shared reward network (e.g., Blood & Zatorre, 2001; Brown, Gao, Tisdelle, Eickhoff, & Liotti, 2011; Koelsch, 2018; Lacey et al., 2011; Mallik, Chanda, & Levitin, 2017; Menon & Levitin, 2005; Montague & Berns, 2002; Sescousse, Caldu, Segura, & Dreher, 2013; Vartanian & Skov, 2014). This network comprises the prefrontal cortex (orbitofrontal, insula, and anterior cingulate cortices) and subcortical limbic structures (nucleus accumbens, ventral pallidum, and amygdala). A mid-anterior subregion of the orbitofrontal cortex has been associated with the subjective feeling of pleasure (Berridge & Kringelbach, 2015). These findings support the view that pleasure induced by artistic stimuli should be studied by comparing behavioral and neuroimaging results across stimulus categories to detect commonalities and differences (Jacobsen & Beudt, 2017; Johnson & Steinerberger, 2019; Marin, 2015; Mas-Herrero, Karhulahti, Marco-Pallarés, Zatorre, & Rodríguez-Fornells, 2018b; Tiihonen et al., 2017; Vessel, Maurer, Denker, & Starr, 2018).

Pleasure and reward are currently understood to involve three different psychological components, namely *liking* (hedonic impact), *wanting* (incentive salience), and *learning* (Pavlovian or instrumental associations and cognitive representations), each of them comprising specific neural networks that interact during a reward–behavior cycle. The neural networks underlying wanting are more largely distributed in the brain than the ones determining liking, which can be limited to several hot spots in the limbic system, but both networks seem generally to work in concert and overlap in the ventral striatum (Berridge & Kringelbach, 2015). All three components can be active at any time in the cycle, but wanting is dominant in the initial appetitive phase (approach behavior), liking in the consummatory phase, and finally, learning in the satiety phase. It is important to note that dopamine is generally not related to the experience of pleasure per se but to the appetitive phase of the reward cycle (Berridge & Kringelbach, 2008; Wise, 2004), regulating motivation and goal-directed behavior. The brainstem is particularly relevant in the discussion of appetitive motivation and motivational arousal associated with dopamine release. Subjective pleasure, in contrast, is primarily based on the release of opioid peptides in the nucleus accumbens (Wise, 2004).

Music, due to its temporal and dynamic nature, is an ideal stimulus to examine different phases of the reward cycle and its underlying neurochemical processes (Chanda & Levitin, 2013). Salimpoor, Benovoy, Larcher, Dagher, and Zatorre (2011) studied dopamine release and its dynamics by using the chill response as an index of pleasure. They found that dopamine release occurs during both the anticipatory and consummatory phases of the reward cycle in two anatomically different subcircuits of the striatum (for causal evidence for the role of dopamine in musical pleasure, see

Ferreri et al., 2019). Motivated by these findings, recent research on the aesthetics of poetry, involving neuroimaging, psychophysiological, and behavioral measures, has revealed that poetry can also induce strong emotions and chills (Wassiliwizky, Koelsch, Wagner, Jacobsen, & Menninghaus, 2017b). Interestingly, the authors only found some overlap between music and poetry regarding the reward-related brain regions, and their results showed in general more posterior activation in the cingulate, insula, and caudate, in comparison with previous reports for music. Differences in activation between music and poetry were also observed when comparing the temporal dynamics of the reward cycle, in which activation of the nucleus accumbens was reported prior, and not during, the emotional peak experience induced by poems.

Studying individual differences has revealed that people who are particularly sensitive to musical reward show an enhanced functional connectivity between the right auditory cortex and the ventral striatum/nucleus accumbens (Salimpoor et al., 2013). Interestingly, the reverse is observed in the case of *musical anhedonia* (Zatorre, 2015), a specific condition in which people do not experience pleasure when listening to music but show intact processing of fundamental pleasures (Martínez-Molina, Mas-Herrero, Rodríguez-Fornells, Zatorre, & Marco-Pallarés, 2016; Mas-Herrero, Zatorre, Rodríguez-Fornells, & Marco-Pallarés, 2014). Using transcranial magnetic stimulation, the same research group has recently provided causal evidence for a link between the activation of the fronto-striatal circuitry with the degree of music-induced pleasure, psychophysiological arousal, and motivational aspects of reward (Mas-Herrero, Dagher, & Zatorre, 2018a).

In the past 15 years, the emerging field of neuroaesthetics (Pearce et al., 2016) has largely advanced our knowledge about the underlying brain processes of aesthetic experiences. In a nutshell, the existing literature emphasizes that attentional, perceptual, affective, and cognitive processes closely interact with each other (Brown et al., 2011; Vartanian & Skov, 2014). For example, during the aesthetic experience of music, brain connectivity networks involving the auditory cortex, the reward system, and brain regions active during mind wandering come into play (Reybrouck, Vuust, & Brattico, 2018; Taruffi, Pehrs, Skouras, & Koelsch, 2017; Wilkins, Hodges, Laurienti, Steen, & Burdette, 2014). During visual aesthetic experiences, activation of the default mode network has also been reported (Belfi et al., 2018; Cela-Conde et al., 2013; Vartanian & Skov, 2014; Vessel, Starr, & Rubin, 2012), highlighting the role of introspective, self-referential thoughts and self-awareness. These findings suggest that there is much more to aesthetic experiences than emotion, pleasure, and reward, and that Berlyne's theory may be too simplistic (see also Pelowski et al., 2017b). Therefore, neuroaesthetics may clearly help determine the nature of aesthetic experiences in the future, for instance, by trying to elucidate whether aesthetic experiences can be characterized as *disinterested pleasure*, as famously proposed by Kant and the third Earl of Shaftesbury in the 18th century. The notion of disinterested pleasure implies that aesthetic experiences are mental states that do not involve the desire to acquire, control, or manipulate the object that induces the experience (we usually tend to want what we like) (Chatterjee & Vartanian, 2016). Chatterjee (2014) has proposed that during aesthetic experiences activation of the liking

system can be observed without activation of the wanting system, but at the moment empirical evidence supporting this hypothesis is scarce (Leder, Gerger, Brieber, & Schwarz, 2014). In summary, the idea that motivational aspects are relevant for the discussion of the nature of aesthetic experiences has also found its place in neuroaesthetics, which in future will certainly employ neuropharmacological methods, enabling causative claims, to a greater extent (for a review, see Spee et al., 2018).

Novelty and Familiarity

Within Berlyne's framework, collative variables are means to either increase or decrease arousal in an individual, and they are tightly linked to information-theoretic concepts. *Novelty* is an arousal-increasing device related to high information content, whereas *familiarity* would be its counterpart; therefore, novelty/familiarity constitute anchors on a continuum. Novelty processing has been associated with the *orienting response* (or "What-is-it?" reflex), which signals changes in the sensory environment (Pavlov, 1927) and is thus essential for survival. At the behavioral level, orienting responses are associated with increased attention toward a stimulus. Repeated exposure to the same stimulus will lead to habituation. At the neural level, a hippocampal–neocortical network (Knight, 1996; Valenti, Mikus, & Klausberger, 2018) is involved in novelty processing and largely, but not only, driven by the dopaminergic system (Rangel-Gomez & Meeter, 2016). Early aestheticians such as Gerard (1764) and Home (1765) considered novelty as one of the main factors affecting aesthetic responses (see Berlyne, 1971), and empirical evidence has supported this view (Berlyne, 1970, 1974b; Leder, 2001; Park, Shimojo, & Shimojo, 2010; Seifert & Chattaraman, 2017; Tinio & Leder, 2009; and see Sluckin, Hargreaves, & Colman, 1983, for a review of older studies).

Berlyne differentiates between *absolute novelty* and *relative novelty*, with the former implying that a person has never interacted with a stimulus before, and the latter that perceived novelty depends on how, for instance, familiar elements appear in a new combination. Moreover, he differentiates between *short-term novelty* and *long-term novelty*, which refers to the time intervals between encountering stimuli or their elements. Examples of short-term novelty refer to the immediate context of elements experienced in aesthetic communication. In the case of music, Berlyne noted (1971) that "a music melody is not only a sequential pattern of pitches but also a sequential pattern of novelty values" (p. 145). In a similar vein, poetic devices such as different rhyme schemes and assonance also generate novelty, familiarity, and repetition through phonological structuring (Berlyne, 1971). Equally important, novelty/familiarity underlies the frequently applied principle of *variation* in music, in which, for example, a melody is repeated and altered by changing some musical features (Sluckin et al., 1983). In the visual domain, this artistic play with familiar and unfamiliar elements can be observed when painters create a series of paintings on one motif, such as Van Gogh's two series of *Sunflowers*, especially in cases where observers have the possibility to experience them

together. The conceptual distinctions of novelty provided by Berlyne are still valid today and underlie current research strands, in which, for instance, memory processes (e.g., recognition) and repeated exposure as well as formed expectations during music and language processing are systematically studied (see also the following section).

Berlyne examined novelty by employing a familiarization paradigm, in which the same stimulus was repeatedly presented; that is, novelty was considered as an independent variable, whose impact on interestingness and hedonic value was measured. He also varied stimulus complexity alongside novelty, and in general, reported for both simple and complex visual stimuli a linear decrease in interestingness with repeated exposure, due to a reduction of uncertainty. For pleasantness ratings, an inverted U-shaped curve was reported only for complex stimuli, probably due to initial stimulus habituation and subsequent satiation (Berlyne, 1970, 1974b). Around the same time, Zajonc (1968) presented his affective model to explain the *mere-exposure effect*, predicting a positive linear relationship between repeated exposure and liking. Contrary to Berlyne, Zajonc argued that the increase in preference is due to the reduction of a negative affective state (i.e., a fear response) initially evoked by a novel stimulus. Since then many researchers have shown interest in the "what is familiar is preferred" phenomenon, which is not only germane to empirical aesthetics but also to fields such as social psychology (Soley & Spelke, 2016) and advertising psychology (Fang, Singh, & Ahluwalia, 2007).

In a very thorough meta-analysis, comparing three prominent theories that may underlie the mere-exposure effect (Berlyne's two-factor theory, fluency theory, and Zajonc's affective model), Montoya, Horton, Vevea, Citkowicz, and Lauber (2017) considered familiarity, recognition, and liking as dependent variables, and further, several moderators, such as age, presentation duration, and presentation type. Based on 268 curve estimates, they found support for an inverted U-shaped curve between exposure frequency and liking for visual, but not for auditory stimuli. In general, Berlyne's model seemed to account for this finding better than fluency theory. However, the authors developed a new theory—the *Representation Matching Model*—because older theories were not able to account for the specific results. This new model, congruent with established models of memory, explains the initial increase in liking with a matching mechanism that compares the incoming stimulus with a formed mental representation (cf. predictive coding in the following section). If the stimulus is perceived as "how it should be" and thus as "correct" and "good", the stimulus is liked. The negative quadratic effect is explained by a habituation via learning mechanism leading to a decrease in liking. In contrast to fluency theory (Reber et al., 2004), the increase in liking is not rooted in the ease of stimulus processing, and the decrease in liking is not explained by fatigue. The authors suggest studying the relationship between recognition, familiarity (as a dependent variable), and liking in greater depth in the future, including different types of stimuli.

Besides studying the mere-exposure effect, novelty/familiarity and its impact on aesthetic responses can be approached from diverse empirical perspectives. For example, researchers have investigated the effect of novelty on cognitive processing and

experience (i.e., a stimulus is perceived as new vs. old; e.g., Bohrn et al., 2013; North & Hargreaves, 1995), and considered familiarity during stimulus selection procedures (Massaro et al., 2012; Nadal et al., 2010) or treated familiarity as a covariate while being primarily interested in other variables (e.g., Forsythe, Mulhern, & Sawey, 2008; Marin, Lampatz, Wandl, & Leder, 2016). This latter aspect is closely related to individual differences in art interest, training, and expertise (Kirk, Skov, Hulme, Christensen, & Zeki, 2009)—variables that are usually discussed in the context of person-related factors. Moreover, novelty/familiarity effects can be investigated at different hierarchical levels of stimulus features. Marin and Bhattacharya (2011) argued that, in the case of music processing, novelty/familiarity effects can be assessed regarding the culture-specific musical system, the musical style, the musical piece, the specific performance of a piece, parts of the musical structure (sections, themes, and motives), and their repetition within a piece, as well as note-by-note effects of musical structure (see the following section). It is a simple task to apply this differentiation to other temporal media, such as literature and dance, and in a slightly modified form, to the study of visual art, in which people may feel familiar or unfamiliar with art periods, styles, artists, as well as with specific artworks and their individual elements. Indeed, theoretical work on modeling aesthetic musical experiences also highlights the role of familiarity (in its different facets), formed mental representations, and spreading activation in associative networks, which may lead to pleasure by a very simple cognitive mechanism (Schubert, Hargreaves, & North, 2014). Hence, different types of novelty/familiarity effects may play a role during various cognitive processing stages of aesthetic stimuli, interacting with each other over time—a hypothesis that appears to be relatively empirically unexplored.

In the domain of music, novelty/familiarity effects are of particular interest, probably due to music's dynamic nature and its inherent repetition of musical elements, a key feature of musical communication (Margulis, 2014). The extant literature on repetition in music, for instance, has studied the detection of musical repetition by untrained listeners (Margulis, 2012), moderating effects of emotional content on psychophysiological (Iwanaga, Kobayashi, & Kawasaki, 2005; Witvliet & Vrana, 2007) and subjective responses (Ali & Peynircioglu, 2010), differences between focused and incidental listening (Schellenberg, Peretz, & Vieillard, 2008) as well as the role of the ecological validity of the stimuli (Szpunar, Schellenberg, & Pliner, 2004). Repetitiveness also seems to increase aesthetic responses to unfamiliar, contemporary art music (Margulis, 2013). It was even shown that repeated exposure to musical excerpts can lead to a transfer of preference to unfamiliar excerpts of the same stylistic genre (Johnston, 2016). Madison and Schiolde (2017) studied familiarity together with musical complexity and demonstrated for four levels of complexity a linear increase in liking ratings after repeated exposure. They also identified familiarity with musical style as a strong predictor of liking (see also Parncutt & Marin, 2006). Janata et al. (2018) manipulated novelty by several musical features and found that instrument entrances are a strong predictor of listening times. In short, this body of research illustrates a wide range of different aspects regarding musical repetition that are currently of interest to the community.

The effects of novelty/familiarity in response to sounds and music have also been investigated using neuroscientific methods. Using a mismatch-negativity (MMN) paradigm, Jacobsen, Schröger, Winkler, and Horváth (2005) showed in an electroencephalography (EEG) study that familiarity with the standard stimuli affected the processing of a deviant sound, as shown by a higher MMN response, in comparison with a condition with a sequence of unfamiliar standard stimuli (cf. predictive coding in the following section). Pereira et al. (2011), using functional magnetic resonance imaging (fMRI), studied "music lovers" with minimal musical training and found that familiar, in contrast to unfamiliar, musical pieces led to increased activation in limbic and paralimbic brain structures as well as in the reward circuit. Van Den Bosch, Salimpoor, and Zatorre (2013) demonstrated that repeated exposure to unfamiliar music increased electrodermal activity (physiological arousal) and felt pleasure. Novel music was able to induce pleasure, but there was no association between emotional arousal and pleasure in this case. Li, Chen, and Tsai (2015) studied repetition in music in relation to different stages of reward processing (reward-anticipation, -gain, and -loss) by focusing on popular songs and the verse–chorus form. Using fMRI they showed that activation of the bilateral temporo-parietal junctions and parts of the orbitofrontal cortex could be related to different stages of the reward cycle. Kumagai, Arvaneh, and Tanaka (2017) conducted an EEG study and were interested in how familiarity with piano music affects cortical entrainment to the musical signal. They found that the brain entrains more strongly to unfamiliar than to familiar music (see also Meltzer et al., 2015). In a thorough meta-analysis, Freitas et al. (2018) compared the neural correlates of familiar and unfamiliar music processing and reported for familiar music activation of the left superior frontal gyrus, the ventral lateral nucleus of the left thalamus, and the left medial surface of the superior frontal gyrus. Unfamiliar music activated the left insula and the right anterior cingulate cortex. Interestingly, limbic structures did not seem to play a role in familiarity processing. These intriguing examples demonstrate that subcortical and cortical brain regions are sensitive to novelty/familiarity during music processing, but at the moment it remains unclear how different facets of novelty/familiarity are represented in the brain and how they interact with each other, which leaves the doors wide open for future experiments.

Surprise and Expectedness

Berlyne (1971) argues that artists can induce arousal changes by playing with expectations, which are linked to the information-theoretic measure of uncertainty and probabilities of events. He states that artists can "violate an expectation by replacing the expected element with one that is distinct from it," they can "introduce a pattern that could be continued in any of several different ways, so that mutually contradictory hypotheses are entertained at once," or they can "introduce a pattern that induces no specifiable expectations at all, so that the appreciator is at complete loss to tell what

might follow. All these possibilities are extensively utilized in all art forms" (1971, p. 145). Berlyne also differentiates between *surprise* and *incongruity*. Surprise can be evoked by elements that follow each other in time, with one element violating the expectations raised by the preceding elements, whereas incongruity may happen by unusual simultaneous combinations of elements. These concepts partly overlap, and during the perception of a painting, for example, there is a time-course of perceived events depending on eye movements, although the elements of a painting occur simultaneously (see Berlyne, 1971). Moreover, it is often the case that what is surprising is novel, and that what is expected is familiar, but as Berlyne pointed out (1971), novelty and surprise are distinct concepts and can occur independently of each other (see also Barto, Mirolli, & Baldassarre, 2013; Manahova, Mostert, Kok, Schoffelen, & de Lange, 2018; Miranda & Ullman, 2007, for some neuroscientific evidence).

Music offers the composer nearly unlimited possibilities to induce pleasure by using unexpected tones, rhythms, timbres, intensity levels, and formal structures—to name only a few options at the level of the musical piece. Haydn's second movement of his symphony No. 94 (popularly known as *Surprise Symphony*) is an example of how music can induce surprise in listeners by presenting a sudden fortissimo chord and then by continuing in a calm manner as if nothing has happened. Another example, at the music-syntactic level, is the *deceptive cadence*, a stylistic means in Western tonal music, in which a violation of the expected chord progression in a cadence (normally indicating the end of a theme or piece) occurs at the final chord, leaving the listener with the unexpected feeling of "oh, it's not over yet." These types of stylistic means are not only inherent to music, but unexpected violations of anticipated events play a role in most types of human works of art.

In literary works and the visual arts, surprise can be induced by violating expectations regarding form *and* semantic content. For instance, the underlying metrical structure of poems and the additional play with deviating stress patterns is a frequent mechanism of generating surprise (or novelty), thus affecting the interpretation of the poem and aesthetic appreciation. At the level of meaning, suspense and surprise-ending stories are commonly found in narrations. Even in static media elements of surprise are a frequent feature of art works, artistic styles, or artists. Some painters of the Renaissance (e.g., Hieronymus Bosch) painted fantastic creatures that were often based on reports of foreign animals, unknown to most people at that time, which probably led to surprise responses. Surprise is also a familiar reaction during the transition period of artistic styles, such as during Impressionism at the end of the 19th century, which violated not only the expected realistic depiction of scenes, but also confronted the audience with unexpected rough-and-ready brush strokes and an unusual use of color. Nowadays, surprise is an essential part of design (Ludden, Schifferstein, & Hekkert, 2008) as well as of performance art in which action artists often address political, social, or religious issues. The ubiquitous use of elements of surprise in various art forms over centuries and their theoretical relevance in philosophical aesthetics (Meyer, 1956b; Morreall, 2016; Skorin-Kapov, 2015, 2016) already hint at its deep root in human cognition (see Berlyne, 1971).

Based on prior experience and learning, expectancies play an essential role in biological functioning because they prepare an organism for reacting to future events and influence—or even determine—perception (de Lange, Heilbron, & Kok, 2018). The anticipation of future events is particularly necessary to avoid dangers, which is directly linked to survival. Due to the inherent uncertainty of future events, the correct prediction of an event is rewarded and induces a positive feeling state because the brain has found an adaptive response in a specific situation (Tooby & Cosmides, 1990). In contrast, surprising events are biologically "bad" and lead to heightened arousal and attention (*action-readiness*). In the context of art perception, however, surprises may initially lead to negative responses that are subsequently re-appraised as nonthreatening and positive (Huron, 2006).

Music is an ideal medium to generate surprise due to its temporal nature. The violation of expectations is one out of several mechanisms that may underlie emotion induction by music (Egermann, Pearce, Wiggins, & McAdams, 2013; Juslin & Västfjäll, 2008; Steinbeis et al., 2006), and strong reactions such as laughter, chills, and awe may be regarded as "flavors of surprise" (Huron, 2006, p. 25). For instance, chill responses, already discussed earlier regarding musical reward processing, are strong emotional responses that can be induced by unexpected changes in acoustic and musical features (Grewe, Nagel, Kopiez, Altenmüller, 2007; Sloboda, 1991). Sudden changes in loudness or harmony or entries of new voices are associated with specific psychophysiological processes including increased arousal (Laeng, Eidet, Sulutvedt, & Panksepp, 2016; Mori & Iwanaga, 2017; Sumpf, Jentschke, & Koelsch, 2015). Besides the possibility to generate surprise by playing with different musical features, surprises are most frequently evoked by violating mental representations of musical syntax (i.e., the structure of music). These cognitive representations are formed during musical enculturation through mechanisms of *implicit* and *statistical learning* (similar to the acquisition of language), which already take place early in life (Saffran, Aslin, & Newport, 1996). Since music is a rule-based system of communication, this enables the generation of schemas and concurrent probabilities of events during music processing (Rohrmeier & Rebuschat, 2012). The hierarchy of musical structures and violations of musical expectations have been intensively studied outside the context of aesthetics for a long time (Krumhansl & Kessler, 1982; Narmour, 1992; Schmuckler, 1989), generating a large body of psychological research on the expectation of melodic tones, chords, meter, and rhythm (Large & Kim, 2019; Rohrmeier & Koelsch, 2012). This research field is mostly characterized by using syntactic priming paradigms, determining the neural basis of expectancy violations, and developing computational models that can predict the probability of the next musical event (Pearce & Rohrmeier, 2018).

In the past two decades, computational models of perception such as *predictive coding* have tried to explain how the brain makes predictions about our environment (de Lange et al., 2018; Rao & Ballard, 1999). The idea that the brain matches sensory input with top-down expectations dates back to Helmholtz (1867), but only now it is possible to make use of computational models and neuroimaging data to test the theory that, for example, perception, decision making, and the processing of language largely depend on

prior knowledge. This model assumes that the world surrounding us is normally stable and thus predictable, and that the brain has built-in internal models about the world (encoded in higher cortical regions) and is matching the incoming sensory input with these models. Prior knowledge not only refers to hierarchical syntactic structures, but also to knowledge about the environment that has been gained over a lifetime (e.g., sunlight comes from above) or to knowledge about the current context. Knowledge can also influence perception on short timescales (e.g., stimuli presented sequentially). If the expectation is fulfilled, this may lead to an increase in perceptual sensitivity for stimulus features and a reduced neural response (*expectation* or *repetition suppression*). On the other hand, an unexpected event may result in a *prediction error* leading to a neural signal. de Lange et al. (2018, p. 766) explain that "the relative impact of expectations versus sensory input on perception depends on their relative reliability (i.e., 'precision'). Observers rely most strongly on prior knowledge when expectations are reliable and stimuli are ambiguous, but rely most strongly on the input when expectations are weak and stimuli reliable." Bayesian probability theory is ideally suited to making such inferences and is used in this context. Importantly, predictive coding models of cortex make predictions about the current sensory input and not forecasts about the next incoming sensory input. This theory has been applied in the context of various types of sensory inputs, including static and dynamic (de Lange et al., 2018).

Predictive coding theory has been adopted to the perception of visual art (Consoli, 2016; Kesner, 2014; Van de Cruys & Wagemans, 2011); for example, suggesting that reward may be generated by an initial state of uncertainty, which turns into a state that is predictable. However, most of this work is of a theoretical nature and empirical evidence is still needed. In the domain of music, predictive coding has recently entered the field by modeling pitch (Chang, Bosnyak, & Trainor, 2018) and rhythm perception (Lumaca, Haumann, Brattico, Grube, & Vuust, 2019; Vuust, Dietz, Witek, & Kringelbach, 2018) as well as by discussing aspects of music-induced reward (Hansen, Dietz, & Vuust, 2017; Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2015). For instance, using three naturalistic pieces of Western art music with different elements of musical surprises, Shany et al. (2019) showed that surprising events led to an increase in subjective arousal and either an increase or decrease in pleasantness compared with unsurprising musical events. At the neuronal level, the authors observed a direct link between surprise-induced pleasantness and increased activation of the nucleus accumbens in a subgroup of participants who enjoyed listening to the music. Moreover, this subgroup also exhibited a stronger connectivity between the nucleus accumbens and auditory cortices related to musical surprises, which demonstrates an association between musical surprises and reward-related activation. In an fMRI study on suspense induced by literary texts, Lehne, Engel, Rohrmeier, Menninghaus, Jacobs, and Koelsch (2015) showed that the emotional experience of suspense was related to activation of regions known from studying social cognition as well as to those related to predictive inference (bilateral activations in inferior frontal regions extending into the lateral premotor cortex in the precentral gyrus). Activation of limbic brain structures was not modulated by suspense. In summary, a review of this current research has revealed that predictive coding has only recently gained

attention in the field, and further, that in the context of dynamic stimuli the term *prediction* is frequently used to refer to anticipation and the prediction of upcoming events (Koelsch, Vuust, & Friston, 2019), which may be conceptually different from studying, for instance, the perception of static visual stimuli or individual musical sounds. In the future, predictive coding in relation to familiarity/novelty (e.g., Vladimirskiy, Urbanczik, & Senn, 2015) may also be studied in greater depth.

Complexity and Simplicity

Humans have to interact with a complex environment, and their sensory system has evolved to deal with this complexity (Godfrey-Smith, 2002). Unsurprisingly then, *complexity/simplicity* is a collative variable that has also gained much attention in the field of empirical aesthetics since the 1950s. The major intuitive determinant of perceived complexity is the number and variety of elements present in a stimulus. For example, a musical piece with several parts played by different instruments is considered as more complex than a piece played by one instrument (Marin & Leder, 2013). Berlyne (1971) argued that in narration, the addition of characters and incidents increases the degree of complexity in a similar way. However, the extensive literature on visual complexity has emphasized that subjective complexity is a multidimensional construct and determined by more than one objective stimulus feature (for a review, see Van Geert & Wagemans, 2020). In fact, the same notion of multidimensionality holds true for musical complexity (Streich, 2007) and literary complexity (Grishakova & Poulaki, 2019), indicating that research on complexity still leaves room for a great diversity of research programs within and across domains.

In the visual domain, complexity and its counterpart simplicity have been linked to concepts such as *goodness of configuration* (Koffka, 1935), information-theoretic concepts of *uncertainty*, *information content*, and *redundancy*, as well as to concepts of *order*, *diversity*, and *structure* (Berlyne, Ogilvie, & Parham, 1968). Berlyne (1960) argued that *subjective complexity* is positively correlated with the number of and dissimilarity between elements and negatively correlated with the perception of how these elements form a unit. Empirical evidence by Berlyne et al. (1968) supported this view by finding two dimensions of subjective complexity, one related to information content determined by "the number of independently selected component elements" (p. 383), and one related to "unitariness vs. articulation into recognizable parts" (p. 384). Whereas Chipman (1977) found that complexity is determined by an interplay between quantitative and structural variables (see also Gartus & Leder, 2017; Ichikawa, 1985; Marin & Leder, 2016), Nadal et al. (2010) suggested three underlying dimensions of subjective complexity, namely (1) the amount and variety of elements, (2) the disorganization of these elements, and (3) asymmetry. These dimensions may show different relationships with hedonic tone. Recent theoretical work by Van Geert and Wagemans (2020) argues that subjective visual complexity stems from *objective complexity* and

order as two separate concepts. The authors define objective complexity as the "quantity and variety of information in a stimulus" (p. 4), similar to the first dimension of Berlyne et al. (1968) and Nadal et al. (2010). Objective order (e.g., symmetry, alignment, iteration) is related to the "structure and organization of the information in the stimulus" (p. 4), which reflects the second dimension of Berlyne et al. (1968) and the second and third dimensions of Nadal et al. (2010). In their framework, simplicity is defined as the opposite of complexity and disorder as the opposite of order. Both complexity and order are considered as multidimensional variables, and Van Geert and Wagemans (2020) suggest that a balance between these two variables may be a good predictor of aesthetic appreciation (see also Eysenck, 1942). Another line of research in the visual domain has focused on fractals, specifically the prevalence and complexity of fractal patterns in visual art (Taylor et al., 2005). Future research will also have to integrate aspects of conceptual and semantic dimensions of complexity (e.g., Commare, Rosenberg, & Leder, 2018; Madan, Bayer, Gamer, Lonsdorf, & Sommer, 2018; Nicki & Moss, 1975) into theories of structurally based complexity and order. Clearly, future studies will thus have to consider different definitions and operationalizations of visual complexity to be able to determine the relationship with aesthetic responses in greater detail.

The development of objective measures of visual complexity has had a long tradition in the field of empirical aesthetics and immensely contributed to our current understanding of subjective complexity. Here, *objective complexity* refers to the complexity that is physically present in a stimulus. Birkhoff (1932) conceived an aesthetic measure (i.e., an aesthetic response) as being determined by two factors, namely order and complexity: M (aesthetic measure) = O (order)/C (complexity). Order (unity or harmony) and complexity (multiplicity and diversity) have been thought to underlie beauty and related measures since the Greek antiquity (Boselie & Leeuwenberg, 1985). Birkhoff was the first to express this relation in a mathematical formula, arguing that order had a positive association with an aesthetic measure and complexity a negative one. However, empirical evidence for Birkhoff 's formula, which was based on the perception of polygons, was scarce, which led Eysenck (1941) to develop a new empirical aesthetic formula, which he also based on studying polygons: $M = O \times C$. In his formula, both order and complexity had a positive contribution to an aesthetic measure, and Eysenck (1968) himself provided empirical evidence for this relationship.

After these initial efforts to formalize aesthetic responses, information-theoretic measures have entered the field, backed up by the growing use of automated algorithms to analyze (and create) artistic materials. Berlyne's *New Experimental Aesthetics* and his followers aimed at elucidating the relationship between objective measures of complexity (uncertainty) and subjective responses. It should be kept in mind that studies linking subjective responses to objective complexity measures have flourished in various fields over the past decades, ranging from basic research on visual aesthetics (for a review, see Van Geert & Wagemans, 2020) to more applied fields such as the aesthetics of food and beverages (Palczak, Blumenthal, Rogeaux, & Delarue, 2019) and the design of websites (Seckler, Opwis, & Tuch, 2015), to name only a few examples. In addition, a considerable amount of research has focused on *how* to measure complexity objectively

in both the visual (Corchs, Ciocca, Bricolo, & Gasparini, 2016; Sun et al., 2018) and auditory domains (Mauch & Levy, 2011; Streich, 2007), proposing a wide range of computational measures detecting structural stimulus features. Needless to say, this endeavor has been largely dependent on computational advances in the field of image statistics and music information retrieval as well as on the development of algorithms and statistical models more generally.

In the field of visual aesthetics and digital image processing, the application of *compression algorithms* has become a widespread measure of complexity (Donderi, 2006; Forsythe et al., 2011; Marin & Leder, 2013). Authors usually report moderate to strong effect sizes regarding the relationship between subjective complexity ratings and image compression ratios, using formats such as *Joint Photographic Expert Group* (JPEG) or *Graphics Interchange Format* (GIF), whose performance depends on the types of images and employed statistical models (Fernandez-Lozano, Carballal, Machado, Santos, & Romero, 2019; Gartus & Leder, 2017; Machado et al., 2015; Marin & Leder, 2013). Compression ratios are an effective measure of visual complexity because, based on *Algorithmic Information Theory*, compression algorithms analyze an image's visual information and try to compress it to the degree that makes a valid reproduction of the input image possible (see also *Kolmogorov complexity*; Kolmogorov, 1968). The file size (in bits) then correlates positively with the complexity of the image. For instance, a simple figure-ground composition can be compressed more than a complex visual scene because simpler images are usually characterized by more redundant information that can be represented by a shorter string of bits (Donderi, 2006).

Other frequently applied measures of objective complexity are *edge detection algorithms*, such as *Perimeter edge detection* and *Canny edge detection* (Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012; Forsythe et al., 2008; Forsythe et al., 2011; Marin & Leder, 2013), and more complex measures such as *root-mean square* (RMS) *contrast* and *phase congruency* (Cavalcante et al., 2014; Gartus & Leder, 2017; Marin & Leder, 2013). These algorithms detect changes in intensity at an image's edges, revealing contours and a global measure of shape, and are sometimes employed in combination with compression rates (Gartus & Leder, 2017). The analysis of the degree of self-similarity constitutes another approach to measure image complexity. For instance, *PHOG self-similarity* is based on the *pyramid histogram of orientation gradients* and compares the similarity of these values for the entire image with the values for individual subparts (Lyssenko, Redies, & Hayn-Leichsenring, 2016; Redies, Amirshahi, Koch, & Denzler, 2012). This measure is closely related to the fractal geometrical properties of an image. Fractals show the same structure when zooming in and out of an image; in other words, there are similar structural patterns at different scales (Mandelbrot, 1967), and the measure of fractal dimension is an indicator of the self-similarity of an image (Theiler, 1990). The aesthetics of fractals has received some attention in the field (Forsythe et al., 2011; Spehar, Clifford, Newell, & Taylor, 2003; Spehar & Taylor, 2013; Taylor, Spehar, Clifford, & Newell, 2008; Viengkham & Spehar, 2018), but the relationship between fractal dimension and subjective aesthetic responses is still unclear. Furthermore, there is a growing body of research on intensity and color-based image

statistics that have shown a relationship with perceived complexity, such as *PHOG anisotropy* (Lyssenko et al., 2016; Redies et al., 2012) and entropy (Marin & Leder, 2013), to name only a few (for a review, see also Van Geert & Wagemans, 2020). Marin & Leder (2016) studied affective environmental scenes and representational paintings, providing initial indications that objective measures may be sensitive to different sub-dimensions of visual complexity, that the subjective multidimensionality of visual complexity only emerges over time, and further, that the relationship between subjective and objective complexity may not depend significantly on the presentation duration of the stimuli. The application of *Convolutional Neural Networks* (CNNs), based on deep learning algorithms, to model visual complexity may also provide one avenue of future research (Krizhevsky, Sutskever, & Hinton, 2012; Simonyan & Zisserman, 2014). Altogether, using a wide range of visual stimuli, future studies will still have to shed more light on which factors determine the relationship between subjective and objective complexity.

Visual complexity has also been studied at the neurophysiological level and event-related potential (ERP) studies have associated visual complexity with early stages of stimulus processing (150 ms), showing effects of the affective image content (i.e., valence) on these early components (Bradley et al., 2007; for a review, see Olofsson, Nordin, Sequeira, & Polich, 2008). However, interaction effects between visual complexity and stimulus valence have also been reported, with complex, as well as more colorful, photographs of objects leading to stronger emotional responses at the neuronal level than simple pictograms (Schlochtermeier et al., 2013). Recently, an *arousal-complexity bias* for picture processing has been documented in a series of experiments, suggesting that arousal may affect the perception of visual complexity (Madan et al., 2018). There is some related neurophysiological evidence in support of this hypothesis (e.g., Schlochtermeier et al., 2013; Schupp, Junghofer, Weike, & Hamm, 2003), which also calls for a further discussion about the conceptual validity of a strict differentiation between emotion and cognition (Ochsner & Phelps, 2007; Okon-Singer, Hendler, Pessoa, & Shackman, 2015). In the field of neuroaesthetics, a recent study using EEG headsets during a museum visit of a large group of participants showed that visual complexity can be predicted above chance-level from EEG signals when viewing aesthetically pleasing art (Kontson et al., 2015). These are promising findings suggesting that combining technological advances with the endeavor to study aesthetic experiences in an ecologically valid setting may well be a fruitful path.

In the musical domain, subjective complexity has also been understood as a multidimensional concept (Conley, 1981; Pressing, 1999; Scheirer, Watson, & Vercoe, 2000; Streich, 2007), depending on a wide range of musical and acoustical features that can be measured objectively. Among the musical features determining general subjective complexity are those related to melody, harmony, rhythm, and meter (only melodic and rhythmic complexity are dealt here with greater detail). Relevant acoustic features of the auditory signal comprise changes in intensity and timbre (Streich, 2007). Numerous researchers have artificially varied one (or two) aspect(s) of musical complexity and investigated its link to hedonic tone, often by integrating objective measures of complexity into their research design, whereas other researchers have studied musical

complexity more holistically with stimuli of higher ecological validity (e.g., Marin & Leder, 2013; North & Hargreaves, 1995). Studies also vary in how much they take into account the dynamic nature of music when modeling complexity by either considering the complexity of a musical excerpt as a whole, the complexity of different parts of an excerpt, or note-to-note relationships within an excerpt. This holds true for the development of cognitive models of musical complexity as well as for approaches aiming to objectively assess musical complexity. Furthermore, recent advances in the development of objective measures build on findings previously reported in the visual domain, such as the application of compression algorithms (e.g., MP3, FLAC, and Ogg Vorbis) to predict subjective musical complexity (de Fleurian, Blackwell, Ben-Tal, & Müllensiefen, 2017; Güçlütürk, Güçlü, van Gerven, & van Lier, 2018; Marin & Leder, 2013). Compression file sizes as measures of objective complexity have also proved to be good predictors when categorizing musical excerpts into musical styles (Cilibrasi, Vitányi, & Wolf, 2004), suggesting practical applications of fundamental research on musical complexity.

Melodic complexity has been widely studied in the field, especially regarding the predictability of note-to-note events (i.e., individual pitches within a sequence of pitches). Melodic complexity depends on music-theoretic factors such as pitch intervals, contours, and hierarchical pitch relationships. It is important to note that there is a conceptual overlap between expectancy violations (Meyer, 1956) and information-theoretic concepts of melodic complexity based on probabilities (Eerola, Himberg, Toiviainen, & Louhivuori, 2006). In simple terms, the less a melody is following our expectations, the more complex it will be perceived. Behind the background of information theory, one can argue that "the less probable a musical utterance is given some baseline probability—for example, the statistical regularities within a corpus—the higher the information entropy of that utterance and the more complex it is likely to be perceived" (Albrecht, 2016, p. 20). Indeed, models of melodic complexity based on either expectancy violations or redundancy appear to perform similarly (Eerola, 2016). Further evidence supporting this view stems from studying musical complexity within the framework of predictive coding: Delplanque, De Loof, Janssens, and Verguts (2019) showed that entropy measures of musical sound sequences (indicating the degree of disorder in a stimulus) may be a good predictor of preference. They authors reported an inverted U-shaped relationship between entropy and preference ratings, which is congruent with Berlyne's psychobiological model.

Drawing on the work of Berlyne and his approach to study both subjective and objective complexity, for example, Crozier (1974) calculated the uncertainty for specific sound sequences and asked how this measure is related to subjective responses to artificial sound-sequence structures in a series of experiments (for a similar approach, see also McMullen, 1974; Vitz, 1966). He reported a linear relationship between subjective complexity and uncertainty level, which was also observed for interestingness, and an inverted U-shaped function between pleasingness (beauty) and uncertainty level, which is in line with Berlyne's psychobiological model (1971). In more recent times, this approach has led to the development of computational models in music research that can not only analyze Shannon entropy as a measure of the uncertainty of the next note

but also the information content of the note that follows as a measure of unexpectedness (Pearce, 2005). This suggests, as indicated above, that collative variables such as complexity and surprise are interwoven in dynamic musical stimuli when generating aesthetic responses (see also Eerola, 2016; Eerola & North, 2000). Melodic complexity has also been examined within the context of fractals (Beauvois, 2007), and an inverted U-shaped curve was also found between aesthetic preference (i.e., melodicity) and subjective and objective complexity measures, respectively. Martins, Gingras, Puig-Waldmueller, and Fitch (2017) also studied musical fractals and their results indicate that recursion and its cognitive presentation may be domain-general.

Recent neurophysiological studies have shown that, for example, event-related potential (ERP) markers, such as the N2, are sensitive to melodic complexity (Minati et al., 2010). In the context of predictive coding, several other studies have found correlates of the brain's precision of its predictive model regarding pitch sequences and their statistical properties (e.g., Garrido, Sahani, & Dolan, 2013; Heilbron & Chait, 2018; Hsu, Le Bars, Hämäläinen, & Waszak, 2015). Using more ecologically valid stimuli than previous studies, Quiroga-Martinez et al. (2019) employed magnetoencephalography to study expectancies within the context of low- and high-entropy melodies varying in their repetitiveness of pitches, which were assessed with a computational model of auditory expectation. Mismatch negativity (MMNm) amplitudes were used as a measure of prediction error, and pitch deviants in a high-entropy context yielded lower MMNm amplitudes than those in a low-entropy context. This finding is in line with theories suggesting that precision-weighted prediction error is a fundamental principle underlying brain function.

Rhythm is defined as a pattern of discrete durations of events that largely depend on perceptual grouping mechanisms (Fraisse, 1963), and that is usually studied by employing measures of human perception and performance (i.e., human reproduction ability, as measured in a tapping task, for instance). Cognitive models of *rhythmic complexity* (for a review, see Honing & Bouwer, 2019) have a long tradition in the field of music psychology (Povel & Essens, 1985; Shmulevich & Povel, 2000), and factors such as tempo and musical expertise have been found to play a role in rhythm perception and production. When we perceive a musical rhythm, we usually perceive a beat, because some musical events are more prominent than others due to accents. Rhythm (beat) perception was reported to activate motor regions of the brain (Grahn & Brett, 2007; Zatorre, Chen, & Penhune, 2007), and rhythmic complexity and musical expertise modulate the synchronization between auditory and motor areas in the brain (Chen, Penhune, & Zatorre, 2008). Musical rhythm has also been discussed within the framework of predictive coding (Vuust & Witek, 2014), and recent evidence suggests that rhythmic sequences of high entropy are more difficult to model for the brain (Lumaca et al., 2019; for a review on the neural underpinnings of rhythm processing, see Nguyen, Gibbings, & Grahn, 2018).

Several studies have also yielded important insights into how rhythmic complexity can be objectively measured, either from an abstract representation or directly from the audio signal. For instance, Alexander and Carey (1968) proposed that subjective

complexity is negatively associated with the number of symmetric subsequences within a rhythmic pattern. This measure of *subsymmetry* was previously successfully applied to visual complexity by the same authors, and for musical rhythms, the measure of subsymmetry was also found to perform better than the *syncopation index* of Longuet-Higgins and Lee (1984) (see also Toussaint & Beltran, 2013, for more recent evidence). Another simple approach to model rhythmic complexity is to measure the rate of rhythmic activity and tempo of rhythms relative to other rhythms (Conley, 1981). de Fleurian et al. (2017) reported that information-theoretic measures based on Kolmogorov complexity and Shannon entropy were predictive of rhythmic complexity, and further, that musical expertise also affected rhythm perception. Other studies revealed that beat-based measures (Povel & Essens, 1985; Shmulevich & Povel, 2000) were good predictors of perceived rhythmic complexity, which also depended on tempo and musical training (Vinke, 2010). Rhythm has also been studied within the context of entrainment and groove (Levitin, Grahn, & London, 2018), and thus it is not surprising that an objective measure of *danceability* based on the audio signal has been proposed (Streich, 2007).

In both the visual and musical domains, individual differences regarding complexity perception have been observed. Berlyne (1971, 1974a) was already aware of individual differences regarding the relationship between complexity and preference (Lane, 1968; Vitz, 1966), and further, that these differences can also be due to how humans respond to ecological and semantic variables. With the advent of modern statistical modeling techniques, such as linear mixed models, which can account for individual differences and integrate them into statistical models, recent psychological research has shown an increased interest in participant-specific response patterns. For example, Güçlütürk, Jacobs, and Lier (2016) demonstrated that the inverted U-shaped function between complexity and liking for abstract computer-generated art can be explained by individual differences in preferences for complexity; namely, by two subgroups with one group preferring simple images and one complex ones. In a similar vein, Spehar, Walker, and Taylor (2016) reported four different subgroups of participants characterized by different preferences for fractal patterns. Street, Forsythe, Reilly, Taylor, and Helmy (2016) considered factors such as gender, age, and culture, and observed main effects of gender and culture as well as an interaction between gender and age on complexity responses to visual fractals. Personality traits can also explain why people respond differently to visual complexity. Marin and Leder (2018) showed that, among females, affect-related personality traits such as trait emotional intelligence, empathy, and stress reactivity modulate complexity and arousal responses to visual and musical stimuli. Güçlütürk and Lier (2019) reported that people can be separated into a subgroup preferring simple songs and one preferring complex songs, with age and gender as possible underlying factors explaining this group difference. Altogether, this current trend of studying individual differences is linked to the question of whether the field can make universal claims about the relationship of core variables such as complexity and liking, or whether a more nuanced way of looking at aesthetic responses is warranted in the future.

Ambiguity and Clarity

Ambiguities are common in everyday life, especially in verbal communication and humor (Martin & Ford, 2018; Rodd, Gaskell, & Marslen Wilson, 2002). According to Berlyne (1971), *ambiguity* and multiple meanings are also central concepts employed by artists and thus highly relevant in the perception of art (Zeki, 2002). The main difference between the perception of ambiguities in everyday life and their perception in relation to art may be that the former is based on prior knowledge and the latter mostly on conventions (Mamassian, 2008). The following examples demonstrate the widespread use of ambiguities in various art forms. In music, a chord can be ambiguous and give rise to different harmonic meanings, which can be resolved by the musical context that follows (Agawu, 1994). For instance, the familiar *Tristan chord* is highly ambiguous and has led to much theoretical speculation regarding its harmonic function in a tonal context. Another example is ambiguity in the analysis of musical form (de Clercq, 2017). In the realm of literature, ambiguity is a commonly found literary device, such as exemplified in *The Sick Rose* by William Blake, which leaves much room for interpretation because of several possible meanings of the words "Rose," "sick," and "worm." In visual arts, for instance, *Cubists* are known for applying ambiguities at several levels, such as depicting different profiles of faces at the same time (e.g., *Seated Woman* by Pablo Picasso) and overlapping shapes of objects with segments belonging to more than one object. Another example is *Op Art*, which is based on visual illusions based on Moiré patterns, figure-ground ambiguities, and three-dimensional appearances based on two-dimensional patterns (Berlyne, 1971, for other examples, see Mamassian, 2008).

Berlyne (1971) refers to Kaplan and Kris (1948), who differentiated between four types of ambiguity that can be discussed in relation to a pattern: (1) *disjunctive ambiguity* (i.e., several alternative and mutual exclusive meanings exist), (2) *additive ambiguity* (i.e., although meanings are mutually exclusive they overlap to a certain extent), (3) *conjunctive ambiguity* (i.e., the two meanings jointly affect the interpretation), and (4) *integrative ambiguity* (i.e., two meanings support each other and evoke one complex meaning). Although much theoretical work has been done on ambiguities (Berndt & Koepnick, 2018; Cross, 2005; Empson, 1930; Gamboni, 2002), empirical research on the role of ambiguity is rather limited in its scope and different conceptions of ambiguity are used, which makes it difficult to compare results across studies.

There is a growing body of research on visual ambiguity, in which, for instance, some researchers focus on *perceptual ambiguity* during early stimulus processing, which is sometimes discussed within the framework of processing fluency, and others on *cognitive ambiguity*, which refers to stable percepts (i.e., one visual experience) with more than one possible meaning or interpretation. Jakesch, Leder, and Forster (2013), for instance, investigated in six experiments how processing fluency (manipulated by presentation duration) and ambiguity interact during the perception of surrealist paintings by René Magritte. They reported that ambiguous paintings were preferred to

nonambiguous paintings at presentation durations of 500 and 1,000 ms, and that ambiguous paintings were harder to process in general. In a follow-up study using the same type of stimuli and facial electromyography, Jakesch, Goller, and Leder (2017) demonstrated that ambiguous Magritte paintings elicited more positively valenced emotional responses than unambiguous paintings, although there was no difference in explicit liking ratings, which may have been due to individual differences. Another recent study by the same group (Markey, Jakesch, & Leder, 2019) explored ERP amplitudes in response to semantic and syntactic inconsistencies inherent to Magritte paintings and also to the photographic versions of these paintings. The data indicated that both types of stimuli were similarly processed and that the N400, a frequently reported ERP component regarding semantic violations, and the P600, an ERP component associated with syntactic violations, were not modulated by semantic and syntactic inconsistencies as expected. This may be due to a specific processing mode activated during the perception of these artworks.

Ambiguity in visual art can also be studied in relation to cognitive insight and the dynamics underlying this process. Muth and Carbon (2013) demonstrated that the level of insights into perceptual Gestalt affect liking and may enhance the appreciation of a stimulus after a perceptual insight (see also Muth, Ebert, Markovic, & Carbon, 2019). Muth, Hesslinger, and Carbon (2015) further reported that modern and contemporary ambiguous artworks that evoke cognitive challenges are appreciated and considered as interesting, and further, that the elaboration of an artwork may be an appealing process in itself. The same research group also developed a *Continuous Evaluation Procedure* (CEP) that allows for the dynamic assessment of subjective responses. This method was used to investigate changes in liking and interest while watching an artistic movie depicting the evolution and metamorphosis of an emerging Gestalt until moments of insights were reached. Their findings indicated that insights increase liking and that interest already increases 1,500 ms before the moments of insights. In a follow-up study using the same artistic video, Muth, Raab, and Carbon (2016) framed their research within the context of predictive coding and reward by having introduced the multidimensional concept of *semantic instability* (Muth & Carbon, 2016). Semantic instability comprises four clusters, namely integrative blend, multistability, indeterminacy, and contrast to perceptual habits (Muth, Hesslinger, & Carbon, 2018). Muth et al. (2016) specifically studied the role of the episodic context (increasing vs. decreasing visual determinacy) in the relationship between determinacy and liking, reporting that in the group that was exposed to an increase of visual determinacy of hidden Gestalts (i.e., the first episode was the least evocative of a clear Gestalt and had high semantic instability) determinacy was a good predictor of liking and that *Aesthetic Aha* effects (i.e., an increase in liking at moments of insight) were stronger in this condition than in the one of decreasing visual determinacy. In other words, the increase in certainty was probably more rewarding in the condition of increasing visual determinacy. The same research group also investigated whether experiencing the artistic video in a laboratory context vs. in an art gallery would affect determinacy and liking. Gallery visitors appreciated the video more and considered it as less semantically unstable than participants in the

laboratory (Muth, Raab, & Carbon, 2017). In short, this research is a valuable contribution to our understanding of how aesthetic experiences may change over time, which also motivates a very naturalistic way of studying art perception by using advanced technological and statistical methods.

Empirical studies on ambiguities have also found its place outside the vision science community. A recent qualitative study by Maksimainen, Eerola, and Saarikallio (2019) examined *emotional ambiguities* present in the experience of either musical or visual objects. Participants were asked to think about an object that induced pleasure, and which was considered as meaningful in their daily lives. They freely reported on any emotions and emotional contradictions during such an experience, rated the strength of the felt pleasure and intensity of emotions as well as to what degree the experience could be regard as a bodily and aesthetic experience. The findings indicated that the majority of participants (~70%) experienced some degree of emotional ambivalence and that this was more frequently the case in response to musical than visual objects. Around half of the emotion terms (46%) used to describe ambivalence were positive, 36% negative, and around 14% both positive and negative. The findings also indicated that ambivalent emotions were associated with slightly higher emotional intensity and considered as aesthetic experiences. The authors also identified different categories of ambivalent attitude types inherent in their descriptions, demonstrating the need to integrate social, ideological, and contextual aspects in the discussion of ambivalent emotions.

One other line of research has tried to disambiguate the meaning and appreciation of paintings (or musical pieces) by offering various types of additional information at the cognitive level, such as the title of the paintings (musical pieces), the name of the painter (composer or performer), or statements about the paintings (music) (e.g., Angladа-Tort, Steffens, & Müllensiefen, 2019; Jakesch & Leder, 2009; Leder, Carbon, & Ripsas, 2006; Margulis, Levine, Simchy-Gross, & Kroger, 2017; Millis, 2001). Although these studies show interesting effects on aesthetic appreciation, they do not deal directly with the ambiguity of the actual stimulus characteristics and are thus not considered here in greater detail.

Major Challenges, Goals, and Suggestions

The current review of the literature on the role of collative variables in aesthetic experiences has shown that Berlyne's legacy is still strongly present among researchers in the field. Collative variables of different stimulus categories and sensory domains are widely studied, using behavioral, computational, and neuroscientific methods. However, Berlyne's theory of collative motivation has not been further developed and his conception of arousal systems is outdated. Notably, trying to replicate the inverted U-shaped curve between a collative variable and hedonic tone is a frequent endeavor,

which has sometimes, but not always, failed. Thus, several major challenges and goals for researchers can be identified. First, if researchers in the field agree that collative variables play a role in aesthetic experiences and that Berlyne's theory is outdated, one challenge will be to integrate these variables into other, more recent and complex models of aesthetic experiences and to determine their relative role. Here I am not only referring to the relative role of collative variables among a set of a collative variables, but also to the relative role of collative variables with respect to other stimulus features (e.g., balance, color, and texture), person-related factors, and the (social) context of aesthetic experience. Second, current research has repeatedly demonstrated that individual differences in responses to collative variables do exist, as already indicated by Berlyne (1971). It will be a challenge to determine the factors underlying these individual differences, and more generally, to integrate individual differences to a larger extent into models of aesthetic experiences. Third, another recent trend is to compare aesthetic responses to collative variables across sensory or artistic domains. Future research will have to clarify whether specific or general models of aesthetic experiences are more appropriate, or in other words, whether we can identify an overlap in addition to domain/stimulus-specific facets at various levels of aesthetic processing and experience. It is also likely that a comparison will reveal differential results for individual collative variables. Fourth, Berlyne's idea to consider motivation in the study of aesthetic responses has recurred in the growing literature on pleasure, the brain, and the reward system. Here, the challenge will be to get a better understanding of the possible role of emotional and/or motivational arousal in relation to reward and pleasure. Fifth, the current review of previous research has revealed that there is much room for replication of reported effects and meta-analyses, considering more nuanced definitions underlying collative variables when generating stimuli for experiments, which may be achieved by acknowledging to a larger degree theoretical work rooted in the humanities. Finally, the field needs to tackle the challenge of studying conscious and nonconscious aspects of aesthetic responses to collative variables by embracing multiple empirical methods. This is especially true for the study of affective (arousal) responses during aesthetic experiences.

Collative variables have been widely studied outside the field of empirical aesthetics, as already mentioned. Examples range from website design to architecture and from music recommendation systems to music and consumer behavior, with frequent references to Berlyne's theory in these applied fields. This demonstrates that the study of collative variables has already found its place outside basic research, which indirectly highlights the relevance of collative variables for aesthetic responses. The literature stemming from more applied fields is growing in parallel to the one stemming from basic research in empirical aesthetics, which researchers working in the respective fields are often unaware of. This hinders fruitful communication across fields, which may be overcome in the future.

One practical suggestion for teaching theories of aesthetic responses at university level would be to avoid reducing Berlyne's theory to the well-known inverted U-shaped curve, as is often done, but to present it instead within the broader context of motivation, to teach its inherent complexity and roots in psychobiology, and finally, to relate

it to the latest findings of neuroimaging research, which have recently put pleasure, reward, and predictive coding into the foreground. Such an approach would do justice to the theoretical contribution of Daniel Berlyne to the field of empirical aesthetics during the past decades, and also show how older theories may re-emerge, and, moreover, how they can be more easily empirically tested thanks to technological and methodological progress.

Acknowledgments

I would like to express my sincere appreciation to Bruno Gingras, Roberto Viviani, Helmut Leder, Martina Jakesch, and Andreas Gartus, whose comments have greatly improved this chapter.

References

Agawu, K. (1994). Ambiguity in tonal music: A preliminary study. In Pople, A. (Ed.), *Theory, analysis and meaning in music* (pp. 86–107). Cambridge, England: Cambridge University Press,

Albrecht, J. (2016). Modeling musical complexity: Commentary on Eerola (2016). *Empirical Musicology Review*, *11*(1), 20–26.

Alexander, C., & Carey, S. (1968). Subsymmetries. *Perception and Psychophysics*, *4*, 73–77.

Ali, S. O., & Peynircioggglu, Z. F. (2010). Intensity of emotions conveyed and elicited by familiar and unfamiliar music. *Music Perception: An Interdisciplinary Journal*, *27*(3), 177–182.

Andretic, R., van Swinderen, B., & Greenspan, R. J. (2005). Dopaminergic modulation of arousal in Drosophila. *Current Biology*, *15*(13), 1165–1175.

Anglada-Tort, M., Steffens, J., & Müllensiefen, D. (2019). Names and titles matter: The impact of linguistic fluency and the affect heuristic on aesthetic and value judgements of music. *Psychology of Aesthetics Creativity and the Arts*, *13*(3), 277–292.

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology*, *4*, 907.

Beauvois, M. W. (2007). Quantifying aesthetic preference and perceived complexity for fractal melodies. *Music Perception*, *24*(3), 247–264.

Belfi, A. M., Vessel, E. A., Brielmann, A., Isik, A. I., Chatterjee, A., Leder, H., … Starr, G. G. (2018). Dynamics of aesthetic experience are reflected in the default-mode network. *NeuroImage*, *188*, 584–597.

Benedek, M., & Kaernbach, C. (2011). Physiological correlates and emotional specificity of human piloerection. *Biological Psychology*, *86*(3), 320–329.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, *45*(3), 180–191.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill.

Berlyne, D. E. (1965). *Structure and direction in thinking*. New York, NY: Wiley.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*(5), 279–286.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

Berlyne, D. E. (1972a). Ends and means of experimental aesthetics. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *26*(4), 303–325.

Berlyne, D. E. (1972b). Humor and its kin. In Goldstein, J. H., & McGhee, P. E. (Eds.), *The psychology of humor: Theoretical perspectives and empirical issues* (pp. 43–60), New York and London: Academic Press.

Berlyne, D. E. (1974a). The New Experimental Aesthetics. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics: Steps towards an objective psychology of aesthetic appreciation* (pp. 1–26). Washington, DC: Hemisphere.

Berlyne, D. E. (1974b). Novelty, complexity, and interestingness. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation* (pp. 175–180). Washington, DC: Hemisphere.

Berlyne, D. E. (1974c). *Studies in the New Experimental Aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Washington, DC: Hemisphere.

Berlyne, D. E., & Lawrence, G. H. (1964). Effects of complexity and incongruity variables on GSR, investigatory behavior, and verbally expressed preference. *The Journal of General Psychology*, *71*(1), 21–45.

Berlyne, D. E., & Madsen, K.B. (1973). *Pleasure, reward, preference: Their nature, determinants, and role in behavior*. Cambridge, England: Academic Press.

Berlyne, D. E., Ogilvie, J., & Parham, L. (1968). The dimensionality of visual complexity, interestingness, and pleasingness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *22*(5), 376.

Berndt, F., & Koepnick, L. (Eds.). (2018). *Ambiguity in contemporary art and theory. Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft*. Hamburg: Felix Meiner Verlag.

Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, *199*(3), 457–480.

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron*, *86*(3), 646–664.

Bestelmeyer, P. E. G., Kotz, S. A., & Belin, P. (2017). Effects of emotional valence and arousal on the voice perception network. *Social Cognitive and Affective Neuroscience*, *12*(8), 1351–1358.

Birkhoff, G. D. (1932). *Aesthetic measure*. Cambridge, MA: Harvard University Press.

Blijlevens, J., Carbon, C. C., Mugge, R., & Schoormans, J. P. (2012). Aesthetic appraisal of product designs: Independent effects of typicality and arousal. *British Journal of Psychology*, *103*(1), 44–57.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823.

Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2013). When we like what we know—A parametric fMRI analysis of beauty and familiarity. *Brain and Language*, *124*(1), 1–8.

Bolles, R. C. (1975). *Theory of motivation*. New York, NY: Harper & Row.

Boselie, F., & Leeuwenberg, E. (1985). Birkhoff revisited: Beauty as a function of effect and means. *The American Journal of Psychology*, *98*(1), 1–39.

Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, *44*(3), 364–373.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59.

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology*, *4*, 206.

Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica*, *154*, 36–42.

Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PLoS ONE*, *9*(6), e99019.

Brown, S. (2019). A unifying model of the arts: The Narration/Coordination Model. *Empirical Studies of the Arts*, *37*(2), 172–196.

Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, *58*(1), 250–258.

Carbon, C. C., & Jakesch, M. (2013). A model for haptic aesthetic processing and its implications for design. *Proceedings of the IEEE*, *101*(9), 2123–2133.

Carretié, L., Hinojosa, J. A., López-Martín, S., & Tapia, M. (2007). An electrophysiological study on the interaction between emotional content and spatial frequency of visual stimuli. *Neuropsychologia*, *45*(6), 1187–1195.

Cavalcante, A., Mansouri, A., Kacha, L., Barros, A. K., Takeuchi, Y., Matsumoto, N., & Ohnishi, N. (2014). Measuring streetscape complexity based on the statistics of local contrast and spatial frequency. *PLoS ONE*, *9*(2), e87097.

Cela-Conde, C. J., García-Prieto, J., Ramasco, J. J., Mirasso, C. R., Bajo, R., Munar, E., Flexas, A., . . . Maestú, F. (2013). Dynamics of brain networks in the aesthetic appreciation. *Proceedings of the National Academy of Sciences*, *110*(Suppl. 2), 10454–10461.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences*, *17*(4), 179–193.

Chang, A., Bosnyak, D. J., & Trainor, L. J. (2018). Beta oscillatory power modulation reflects the predictability of pitch change. *Cortex*, *106*, 248–260.

Chatterjee, A. (2014). *The aesthetic brain: How we evolved to desire beauty and enjoy art*. Oxford, England: Oxford University Press.

Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, *1369*(1), 172–194.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239.

Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, *41*(6), 631–647.

Chipman, S. F. (1977). Complexity and structure in visual patterns. *Journal of Experimental Psychology: General*, *106*(3), 269–301.

Chmiel, A., & Schubert, E. (2017). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music*, *45*(6), 886–909.

Chmiel, A. & Schubert, E. (2018). Emptying rooms: When the inverted-U model of preference fails—An investigation using music with collative extremes. *Empirical Studies of the Arts*, *36*(2), 199–221.

Cilibrasi, R., Vitányi, P., & Wolf, R. D. (2004). Algorithmic clustering of music based on string compression. *Computer Music Journal*, *28*(4), 49–67.

Colver, M. C., & El-Alayli, A. (2016). Getting aesthetic chills from music: The connection between openness to experience and frisson. *Psychology of Music*, *44*(3), 413–427.

Commare, L., Rosenberg, R., & Leder, H. (2018). More than the sum of its parts: Perceiving complexity in painting. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(4), 380–391.

Conley, J. K. (1981). Physical correlates of the judged complexity of music by subjects differing in musical background. *British Journal of Psychology*, *72*(4), 451–464.

Consoli, G. (2016). Predictive error reduction and the twofold nature of aesthetic pleasure. *Art & Perception*, *4*(4), 327–338.

Corchs, S. E., Ciocca, G., Bricolo, E., & Gasparini, F. (2016). Predicting complexity perception of real world images. *PLoS ONE*, *11*(6), e0157986.

Cross, I. (2005). Music and meaning, ambiguity and evolution. In D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.), *Musical communication* (pp. 27–44). New York, NY: Oxford University Press.

Crozier, J. B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics*. *Steps towards an objective psychology of aesthetic appreciation* (pp. 27–90). New York, NY: John Wiley & Sons.

Cui, A. X., Collett, M. J., Troje, N. F., & Cuddy, L. L. (2015). Familiarity and preference for pitch probability profiles. *Cognitive Processing*, *16*(2), 211–218.

Cupchik, G. C. (1988). The legacy of Daniel E. Berlyne. *Empirical Studies of the Arts*, *6*(2), 171–186.

Day, H. (Ed.). (1981). *Advances in intrinsic motivation and aesthetics*. New York, NY: Plenum Press.

de Clercq, T. (2017). Embracing ambiguity in the analysis of form in pop/rock music, 1982–1991. *Music Theory Online*, *23*(3), 1–20.

de Fleurian, R., Blackwell, T., Ben-Tal, O., & Müllensiefen, D. (2017). Information-theoretic measures predict the human judgment of rhythm complexity. *Cognitive Science*, *41*(3), 800–813.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, *22*(9), 764–779.

Delplanque, J., De Loof, E., Janssens, C., & Verguts, T. (2019). The sound of beauty: How complexity determines aesthetic preference. *Acta Psychologica*, *192*, 146–152.

Donderi, D. C. (2006). Visual complexity: A review. *Psychological Bulletin*, *132*(1), 73–97.

Eerola, T. (2016). Expectancy-violation and information-theoretic models of melodic complexity. *Empirical Musicology Review*, *11*(1), 2–17.

Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of Western and African folk melodies by Western and African listeners. *Psychology of Music*, *34*(3), 337–371.

Eerola, T., & North, A. C. (2000). Expectancy-based model of melodic complexity. *Proceedings of the sixth international conference on music perception and cognition*. Keele, England: Department of Psychology. CD-ROM.

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 533–553.

Empson, W. (1930). *Seven types of ambiguity*. London: Random House.

Eysenck, H. J. (1941). The empirical determination of an aesthetic formula. *Psychological Review*, *48*(1), 83–92.

Eysenck, H. J. (1942). The experimental study of the "good Gestalt"—a new approach. *Psychological Review*, *49*(4), 344–364.

Eysenck, H. J. (1968). An experimental study of aesthetic preference for polygonal figures. *The Journal of General Psychology*, *79*(1), 3–17.

Fang, X., Singh, S., & Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of Consumer Research*, *34*(1), 97–103.

Fechner, G. T. (1876). *Vorschule der Ästhetik*. Leipzig, Germany: Breitkopf & Härtel.

Fernandez-Lozano, C., Carballal, A., Machado, P., Santos, A., & Romero, J. (2019). Visual complexity modelling based on image features fusion of multiple kernels. *PeerJ*, *7*, 28.

Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripollés, P., Gomez-Andres, A., Alicart, H., . . . Rodriguez-Fornells, A. (2019). Dopamine modulates the reward experiences elicited by music. *Proceedings of the National Academy of Sciences*, *116*(9), 3793–3798.

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, *40*(1), 116–129.

Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, *102*, 49–70.

Fraisse, P. (1963). *The psychology of time*. New York, NY: Harper & Row.

Freitas, C., Manzato, E., Burini, A., Taylor, M. J., Lerch, J. P., & Anagnostou, E. (2018). Neural correlates of familiarity in music listening: A systematic review and a neuroimaging meta-analysis. *Frontiers in Neuroscience*, *12*, 686.

Gamboni, D. (2002). *Potential images. Ambiguity and indeterminancy in modern art*. London: Reaction Books Ltd.

Garrido, M. I., Sahani, M., & Dolan, R. J. (2013). Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS Computational Biology*, *9*(3), e1002999.

Gartus, A., & Leder, H. (2017). Predicting perceived visual complexity of abstract patterns using computational measures: The influence of mirror symmetry on complexity perception. *PLoS ONE*, *12*(11), e0185276.

Gerard, A. (1764). *An essay on taste*. London: Millar.

Godfrey-Smith, P. (2002). Environmental complexity and the evolution of cognition. In R. J. Sternberg & J. C. Kaufman (Eds.): *The evolution of intelligence* (pp. 233–249). New Jersey: Lawrence Erlbaum Associates Publishers.

Graf, L. K. M., & Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: The pleasure-interest model of aesthetic liking. *Personality and Social Psychology Review*, *19*(4), 395–410.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception: An Interdisciplinary Journal*, *24*(3), 297–314.

Grishakova, M., & Poulaki, M. (2019). *Narrative complexity: Cognition, embodiment, evolution*. Lincoln, NB: University of Nebraska Press.

Güçlütürk, Y., Güçlü, U., van Gerven, M., & van Lier, R. (2018). Representations of naturalistic stimulus complexity in early and associative visual and auditory cortices. *Scientific Reports*, *8*(1), 3439.

Güçlütürk, Y., Jacobs, R. H. A. H., & Lier, R. V. (2016). Liking versus complexity: Decomposing the inverted U-curve. *Frontiers in Human Neuroscience*, *10*, 112.

Güçlütürk, Y., & Lier, R. V. (2019). Decomposing complexity preferences for music. *Frontiers in Psychology*, *10*, 674.

Hansen, N. C., Dietz, M. J., & Vuust, P. (2017). Commentary: Predictions and the brain: How musical sounds become rewarding. *Frontiers in Human Neuroscience*, *11*, 168.

Heilbron, M., & Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience*, *389*, 54–73.

Hekkert, P., & Leder, H. (2008). Product aesthetics. In H. N. J. Schifferstein, & P. Hekkert (Eds.), *Product experience* (pp. 259–285). San Diego, CA: Elsevier.

Heyduk, R. G. (1975). Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception & Psychophysics*, *17*(1), 84–90.

Home, H. (1765). *Elements of criticism*. London: Millar.

Honing, H., & Bouwer, F. L. (2019). Rhythm. In P. J. Rentfrow, & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 33–69). Cambridge, MA: MIT Press.

Hsu, Y.-F., Le Bars, S., Hämäläinen, J. A., & Waszak, F. (2015). Distinctive representation of mispredicted and unpredicted prediction errors in human electroencephalography. *Journal of Neuroscience*, *35*(43), 14653–14660.

Hull, C. L. (1943). *Principles of behaviour*. New York, NY: Appleton-Century-Crofts.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Ichikawa, S. (1985). Quantitative and structural factors in the judgment of pattern complexity. *Perception & Psychophysics*, *38*(2), 101–109.

Iwanaga, M., Kobayashi, A., & Kawasaki, C. (2005). Heart rate variability with repetitive exposure to music. *Biological Psychology*, *70*(1), 61–66.

Jacobs, A. M., & Kinder, A. (2018). What makes a metaphor literary? Answers from two computational studies. *Metaphor and Symbol*, *33*(2), 85–100.

Jacobsen, T., & Beudt, S. (2017). Domain generality and domain specificity in aesthetic appreciation. *New Ideas in Psychology*, *47*, 97–102.

Jacobsen, T., Schröger, E., Winkler, I., & Horváth, J. (2005). Familiarity affects the processing of task-irrelevant auditory deviance. *Journal of Cognitive Neuroscience*, *17*(11), 1704–1713.

Jakesch, M., Goller, J., & Leder, H. (2017). Positive fEMG patterns with ambiguity in paintings. *Frontiers in Psychology*, *8*, 785.

Jakesch, M., & Leder, H. (2009). Finding meaning in art: Preferred levels of ambiguity in art appreciation. *Quarterly Journal of Experimental Psychology*, *62*(11), 2105–2112.

Jakesch, M., & Leder, H. (2015). The qualitative side of complexity: Testing effects of ambiguity on complexity judgments. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 200–205.

Jakesch, M., Leder, H., & Forster, M. (2013). Image ambiguity and fluency. *PLoS ONE*, *8*(9), e74084.

Janata, P., Peterson, J., Ngan, C., Keum, B., Whiteside, H., & Ran, S. (2018). Psychological and musical factors underlying engagement with unfamiliar music. *Music Perception*, *36*(2), 175–200.

Johnson, S. G. B., & Steinerberger, S. (2019). Intuitions about mathematical beauty: A case study in the aesthetic experience of ideas. *Cognition*, *189*, 242–259.

Johnston, R. R. (2016). The effect of repetition on preference ratings for select unfamiliar musical examples: Does preference transfer? *Psychology of Music*, *44*(3), 514–526.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575; discussion 575–621.

Kaplan, A., & Kris, E. (1948). Esthetic ambiguity. *Philosophy and Phenomenological Research*, *8*(3), 415–435.

Kesner, L. (2014). The predictive mind and the experience of visual art work. *Frontiers in Psychology*, *5*, 1417.

Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *NeuroImage*, *44*(3), 1125–1132.

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, *383*, 256–259.

Koelsch, S. (2009). Neural substrates of processing syntax and semantics in music. In R. Haas, & V. Brandes (Eds.), *Music that works: Contributions of biology, neurophysiology, psychology, sociology, medicine and musicology* (pp. 143–153). Vienna: Springer Vienna.

Koelsch, S. (2018). Investigating the neural encoding of emotion with music. *Neuron*, *98*(6), 1075–1079.

Koelsch, S., Jacobs, A. M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C., & Gebauer, G. (2015). The quartet theory of human emotions: An integrative and neurofunctional model. *Physics of Life Reviews*, *13*, 1–27.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*(3), 302–307.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63–77.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace.

Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics*, *2*(1–4), 157–168.

Konecni, V. J. (2011). The influence of affect on music choice. In P. N. Juslin, & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 697–723). Oxford, England: Oxford University Press.

Konečni, V. J. (1978). Daniel E. Berlyne: 1924–1976. *American Journal of Psychology*, *91*, 133–137.

Konečni, V. J. (1996). Daniel E. Berlyne (1924–1976): Two decades later. *Empirical Studies of the Arts*, *14*(2), 129–142.

Kontson, K., Megjhani, M., Brantley, J., Cruz-Garza, J., Nakagome, S., Robleto, D., . . . Contreras-Vidal, J. (2015). "Your Brain on Art": Emergent cortical dynamics during aesthetic experiences. *Frontiers in Human Neuroscience*, *9*, 626.

Kragel, P. A. & LaBar, K. S. (2016). Decoding the nature of emotion in the Brain. *Trends in Cognitive Sciences*, *20*(6), 444–455.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In P. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Proceedings of the 25th international conference on neural information processing systems-volume 1 (NIPS'12)* (pp. 1097–1105). Red Hook, NY: Curran Associates Inc.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368.

Kumagai, Y., Arvaneh, M., & Tanaka, T. (2017). Familiarity affects entrainment of EEG in music listening. *Frontiers in Human Neuroscience*, *11*, 8.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940.

Lacey, S., Hagtvedt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., … Sathian, K. (2011). Art for reward's sake: Visual art recruits the ventral striatum. *Neuroimage*, *55*(1), 420–433.

Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition*, *44*, 161–178.

Lane, S. H. (1968). Preference for complexity as a function of schematic orderliness and redundancy. *Psychonomic Science*, *13*(4), 209–210.

Large, E. W., & Kim, J. C. (2019). Musical expectancy. In P. J. Rentfrow, & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 221–263). Cambridge, MA: MIT Press.

Leder, H. (2001). Determinants of preference: When do we like what we know? *Empirical Studies of the Arts*, *19*(2), 201–211.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508.

Leder, H., Carbon, C. C., & Ripsas, A. L. (2006). Entitling art: Influence of title information on understanding and appreciation of paintings. *Acta Psychologica*, *121*(2), 176–198.

Leder, H., Gerger, G., Brieber, D., & Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cognition and Emotion*, *28*(6), 1137–1147.

Leder, H., Ring, A., & Dressler, S. G. (2013). See me, feel me! Aesthetic evaluations of art portraits. *Psychology of Aesthetics Creativity and the Arts*, *7*(4), 358–369.

Lee, S.-H., & Dan, Y. (2012). Neuromodulation of brain states. *Neuron*, *76*(1), 209–222.

Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A. M., & Koelsch, S. (2015). Reading a suspenseful literary text activates brain areas related to social cognition and predictive inference. *PLoS ONE*, *10*(5), e0124550.

Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. *Annual Review of Psychology*, *69*(1), 51–75.

Li, C. W., Chen, J. H., & Tsai, C. G. (2015). Listening to music in a risk-reward context: The roles of the temporoparietal junction and the orbitofrontal/insular cortices in reward-anticipation, reward-gain, and reward-loss. *Brain Research*, *1629*, 160–170.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception: An Interdisciplinary Journal*, *1*(4), 424–441.

Low, A., Bradley, M. M., & Lang, P. J. (2013). Perceptual processing of natural scenes at rapid rates: Effects of complexity, content, and emotional arousal. *Cognitive Affective & Behavioral Neuroscience*, *13*(4), 860–868.

Ludden, G. D., Schifferstein, H. N., & Hekkert, P. (2008). Surprise as a design strategy. *Design Issues*, *24*(2), 28–38.

Lumaca, M., Haumann, N. T., Brattico, E., Grube, M., & Vuust, P. (2019). Weighting of neural prediction error by rhythmic complexity: A predictive coding account using mismatch negativity. *European Journal of Neuroscience*, *49*(12), 1597–1609.

Lyssenko, N., Redies, C., & Hayn-Leichsenring, G. U. (2016). Evaluating abstract art: Relation between term usage, subjective ratings, image properties and personality traits. *Frontiers in Psychology*, *7*, 973.

Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta Psychologica*, *160*, 43–57.

Machotka, P. (1980). Daniel Berlyne's contributions to empirical aesthetics. *Motivation and Emotion*, *4*(2), 113–121.

Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2018). Visual complexity and affect: Ratings reflect more than meets the eye. *Frontiers in Psychology*, *8*, 2368.

Madison, G., & Schiolde, G. (2017). Repeated listening increases the liking for music regardless of its complexity: Implications for the appreciation and aesthetics of music. *Frontiers in Neuroscience*, *11*, 147.

Maksimainen, J. P., Eerola, T., & Saarikallio, S. H. (2019). Ambivalent emotional experiences of everyday visual and musical objects. *SAGE Open*, *9*(3).

Mallik, A., Chanda, M. L., & Levitin, D. J. (2017). Anhedonia to music and mu-opioids: Evidence from the administration of naltrexone. *Scientific Reports*, *7*, 41952.

Mamassian, P. (2008). Ambiguities and conventions in the perception of visual art. *Vision Research*, *48*(20), 2143–2153.

Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J.-M., & de Lange, F. P. (2018). Stimulus familiarity and expectation jointly modulate neural activity in the visual ventral stream. *Journal of Cognitive Neuroscience*, *30*(9), 1366–1377.

Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, *156*(3775), 636–638.

Margulis, E. H. (2012). Musical repetition detection across multiple exposures. *Music Perception: An Interdisciplinary Journal*, *29*(4), 377–385.

Margulis, E. H. (2013). Aesthetic responses to repetition in unfamiliar music. *Empirical Studies of the Arts*, *31*(1), 45–57.

Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford, England: Oxford University Press.

Margulis, E. H., Levine, W. H., Simchy-Gross, R., & Kroger, C. (2017). Expressive intent, ambiguity, and aesthetic experiences of music and poetry. *PLoS ONE*, *12*(7), e0179145.

Marin, M. M. (2015). Crossing boundaries: Toward a general model of neuroaesthetics. *Frontiers in Human Neuroscience*, *9*, 443.

Marin, M. M., & Bhattacharya, J. (2011). Music induced emotions: Some current issues and cross-modal comparisons. In J. Hermida, & M. Ferrero (Eds.), *Music education* (pp. 1–38). Hauppauge, NY: Nova Science Publishers.

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience*, *10*, 536.

Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS ONE*, *8*(8), e72412.

Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica*, *163*, 38–58.

Marin, M. M., & Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Personality and Individual Differences*, *125*, 80–90.

Markey, P. S., Jakesch, M., & Leder, H. (2019). Art looks different—Semantic and syntactic processing of paintings and associated neurophysiological brain responses. *Brain and Cognition*, *134*, 58–66.

Martin, R. A., & Ford, T. E. (2018). *The psychology of humor: An integrative approach*. London: Academic Press.

Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *The Journal of Mind and Behavior*, *5*(1), 49–80.

Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 661–670.

Martindale, C., Moore, K., & Anderson, K. (2005). The effect of extraneous stimulation on aesthetic preference. *Empirical Studies of the Arts*, *23*(2), 83–91.

Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J., & Marco-Pallarés, J. (2016). Neural correlates of specific musical anhedonia. *Proceedings of the National Academy of Sciences*, *113*(46), E7337–E7345.

Martins, M. D., Gingras, B., Puig-Waldmueller, E., & Fitch, W. T. (2017). Cognitive representation of "musical fractals": Processing hierarchy and recursion in the auditory domain. *Cognition*, *161*, 31–45.

Mas-Herrero, E., Dagher, A., & Zatorre, R. J. (2018a). Modulating musical reward sensitivity up and down with transcranial magnetic stimulation. *Nature Human Behaviour*, *2*(1), 27–32.

Mas-Herrero, E., Karhulahti, M., Marco-Pallares, J., Zatorre, R. J., & Rodriguez-Fornells, A. (2018b). The impact of visual art and emotional sounds in specific musical anhedonia. In J. F. Christensen, & A. Gomila (Eds.), *Arts and the brain: Psychology and physiology beyond pleasure* (pp. 399–413). Amsterdam: Elsevier Science.

Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology*, *24*(6), 699–704.

Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When art moves the eyes: A behavioral and eye-tracking study. *PLoS ONE*, *7*(5), e37285.

Mauch, M., & Levy, M. (2011). Structural change on multiple time scales as a correlate of musical complexity. In *Proceedings of the 12th international conference on music information retrieval (ISMIR 2011)*. http://matthiasmauch.net/_pdf/mauch_scm_2011.pdf

McMullen, P. T. (1974). Influence of number of different pitches and melodic redundancy on preference responses. *Journal of Research in Music Education*, *22*(3), 198–204.

Meltzer, B., Reichenbach, C. S., Braiman, C., Schiff, N. D., Hudspeth, A. J., & Reichenbach, T. (2015). The steady-state response of the cerebral cortex to the beat of music reflects both the comprehension of music and attention. *Frontiers in Human Neuroscience*, *9*, 436.

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PLoS ONE*, *10*(6), e0128451.

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage*, *28*(1), 175–184.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Millis, K. (2001). Making meaning brings pleasure: The influence of titles on aesthetic experiences. *Emotion*, *1*(3), 320–329.

Minati, L., Salvatoni, L., Rosazza, C., Pietrocini, E., Visani, E., Panzica, F., . . . Franceschetti, S. (2010). Event-related potential (ERP) markers of melodic processing: The N2 component is modulated by structural complexity, not by melodic "meaningfulness." *Brain Research Bulletin*, *83*(1–2), 23–28.

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *NeuroImage*, *38*(2), 331–345.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, *36*(2), 265–284.

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., & Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin*, *143*(5), 459–498.

Mori, K., & Iwanaga, M. (2017). Two types of peak emotional responses to music: The psychophysiology of chills and tears. *Scientific Reports*, *7*, 46063.

Morreall, J. (2016). Philosophy of humor. In E. N. Zalta (ed.), *The Stanford encyclopedia of philosophy* (Winter 2016 edition). Stanford, CA: Stanford University. https://plato.stanford.edu/archives/win2016/entries/humor/.

Moruzzi, G., & Magoun, H. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *1*(1–4), 455–473.

Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, *68*(10), 689–709.

Muth, C., & Carbon, C.-C. (2013). The Aesthetic Aha: On the pleasure of having insights into Gestalt. *Acta Psychologica*, *144*(1), 25–30.

Muth, C., & Carbon, C.-C. (2016). SeIns: Semantic Instability in Art. *Art & Perception*, *4*(1–2), 145–184.

Muth, C., Ebert, S., Markovic, S., & Carbon, C.-C. (2019). "Aha"ptics: Enjoying an Aesthetic Aha during haptic exploration. *Perception*, *48*(1), 3–25.

Muth, C., Hesslinger, V. M., & Carbon, C.-C. (2015). The appeal of challenge in the perception of art: How ambiguity, solvability of ambiguity, and the opportunity for insight affect appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 206–216.

Muth, C., Hesslinger, V. M., & Carbon, C.-C. (2018). Variants of semantic instability (SeIns) in the arts: A classification study based on experiential reports. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(1), 11–23.

Muth, C., Raab, M. H., & Carbon, C.-C. (2016). Semantic stability is more pleasurable in unstable episodic contexts. On the relevance of perceptual challenge in art appreciation. *Frontiers in Human Neuroscience*, *10*, 43.

Muth, C., Raab, M. H., & Carbon, C.-C. (2017). Expecting the unexpected: How gallery visitors experience semantic instability in art. *Art & Perception*, *5*(2), 121–142.

Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts*, *28*(2), 173–191.

Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication-realization model*. Chicago, IL: University of Chicago Press.

Nguyen, T., Gibbings, A., & Grahn, J. (2018). Rhythm and beat perception. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 507–521). Berlin, Heidelberg, Germany: Springer.

Nicki, R., & Moss, V. (1975). Preference for non-representational art as a function of various measures of complexity. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *29*(3), 237–249.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology: A Journal of Research in Music Cognition*, *14*(1–2), 77–93.

North, A. C., & Hargreaves, D. J. (2000). Collative variables versus prototypicality. *Empirical Studies of the Arts*, *18*(1), 13–17.

Ochsner, K. N., & Phelps, E. (2007). Emerging perspectives on emotion-cognition interactions. *Trends in Cognitive Sciences*, *11*(8), 317–318.

Okon-Singer, H., Hendler, T., Pessoa, L., & Shackman, A. J. (2015). The neurobiology of emotion–cognition interactions: Fundamental questions and strategies for future research. *Frontiers in Human Neuroscience*, *9*(58).

Olds, J., & Olds, M. (1965). Drives, rewards, and the brain. *New Directions in Psychology*, *2*, 327–410.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247–265.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Chicago, IL: University of Illinois Press.

Palczak, J., Blumenthal, D., Rogeaux, M., & Delarue, J. (2019). Sensory complexity and its influence on hedonic responses: A systematic review of applications in food and beverages. *Food Quality and Preference*, *71*, 66–75.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*, 77–107.

Park, J., Shimojo, E., & Shimojo, S. (2010). Roles of familiarity and novelty in visual preference judgments are segregated across object categories. *Proceedings of the National Academy of Sciences*, *107*(33), 14552–14555.

Parncutt, R., & Marin, M. M. (2006). Emotions and associations evoked by unfamiliar music. In H. Gottesdiener, & J. C. Vilatte (Eds.), *Culture and communication: Proceedings of the XIXth conference of the international association of empirical aesthetics* (pp. 725–729). Avignon, France: IAEA.

Pavlov, I. P. (1927). *Conditional reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford, England: Oxford University Press.

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition*. Unpublished doctoral thesis, City University London.

Pearce, M. T., & Rohrmeier, M. A. (2018). Musical syntax II: Empirical perspectives. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 487–505). Berlin: Springer-Verlag.

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, *11*(2), 265–279.

Pelowski, M., Forster, M., Tinio, P. P., Scholl, M., & Leder, H. (2017a). Beyond the lab: An examination of key factors influencing interaction with "real" and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(3), 245–264.

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017b). Move me, astonish me . . . delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, *21*, 80–125.

Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in Human Neuroscience*, *10*, 160.

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLoS ONE*, *6*(11), 9.

Picard, R. W., Fedor, S., & Ayzenberg, Y. (2016). Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, *8*(1), 62–75.

Post, R., Nguyen, T., & Hekkert, P. (2017). Unity in variety in website aesthetics: A systematic inquiry. *International Journal of Human-Computer Studies*, *103*, 48–62.

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception: An Interdisciplinary Journal*, *2*(4), 411–440.

Pressing, J. (1999). Complexity and the structure of musical patterns. *Noetica*, *3*(8), 1–8.

Quiroga-Martinez, D. R., Hansen, N. C., Højlund, A., Pearce, M. T., Brattico, E., & Vuust, P. (2019). Reduced prediction error responses in high-as compared to low-uncertainty musical contexts. *Cortex*, *120*, 181–200.

Rangel-Gomez, M., & Meeter, M. (2016). Neurotransmitters and novelty: A systematic review. *Journal of Psychopharmacology*, *30*(1), 3–12.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364–382.

Redies, C., Amirshahi, S. A., Koch, M., & Denzler, J. (2012). PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In A. Fusiella, V. Murino, & R. Cucchiara (Eds.), *Lecture notes in computer science, Vol. 7583* (pp. 522–531), Berlin, Heidelberg, Germany: Springer.

Reybrouck, M., Vuust, P., & Brattico, E. (2018). Brain connectivity networks and the aesthetic experience of music. *Brain Sciences*, *8*(6), E107.

Robbins, T. W., Granon, S., Muir J, L., Durantou, F., Harrison, A., & Everitt B, J. (2006). Neural systems underlying arousal and attention: Implications for drug abuse. *Annals of the New York Academy of Sciences*, *846*(1), 222–237.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*(2), 245–266.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–175.

Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science*, *4*(4), 525–553.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saklofske, D. H. (1975). Visual aesthetic complexity, attractiveness and diversive exploration. *Perceptual and Motor Skills*, *41*(3), 813–814.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*, 257–262.

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE*, *4*(10), e7487.

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, *340*(6129), 216–219.

Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: How musical sounds become rewarding. *Trends in Cognitive Sciences*, *19*(2), 86–91.

Schäfer, T., & Sedlmeier, P. (2010). What makes us like music? Determinants of music preference. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 223–234.

Schäfer, T., & Sedlmeier, P. (2011). Does the body move the soul? The impact of arousal on music preference. *Music Perception: An Interdisciplinary Journal*, *29*(1), 37–50.

Scheirer, E. D., Watson, R. B., & Vercoe, B. L. (2000). On the perceived complexity of short musical segments. In C. Woods (Ed.), *Proceedings of the 2000 international conference on music perception and cognition*. Keele, England: University of Keele.

Schellenberg, E. G., Peretz, I., & Vieillard, S. (2008). Liking for happy- and sad-sounding music: Effects of exposure. *Cognition and Emotion*, *22*(2), 218–237.

Schlochtermeier, L. H., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., & Jacobs, A. M. (2013). Emotional picture and word processing: An fMRI study on effects of stimulus complexity. *PLoS ONE*, *8*(2), e55619.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception: An Interdisciplinary Journal*, *7*(2), 109–149.

Schubert, E. (2010). Affective, evaluative, and collative responses to hated and loved music. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(1), 36–46.

Schubert, E., Hargreaves, D., & North, A. (2014). A dynamically minimalist cognitive explanation of musical preference: Is familiarity everything? *Frontiers in Psychology*, *5*, 38.

Schupp, H. T., Junghofer, M., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1), 7–13.

Seckler, M., Opwis, K., & Tuch, A. N. (2015). Linking objective design factors with subjective aesthetics: An experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. *Computers in Human Behavior*, *49*, 375–389.

Seifert, C. & Chattaraman, V. (2017). Too new or too complex? Why consumers' aesthetic sensitivity matters in apparel design evaluation. *Journal of Fashion Marketing and Management*, *21*(2), 262–276.

Sescousse, G., Caldu, X., Segura, B., & Dreher, J. C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(4), 681–696.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal*, *27*, 379–423, 623–656.

Shany, O., Hendler, T., Singer, N., Tarrasch, R., Gold, B. P., Jacoby, N., & Granot, R. (2019). Surprise-related activation in the nucleus accumbens interacts with music-induced pleasantness. *Social, Cognitive and Affective Neuroscience*, *14*(4), 459–470.

Shmulevich, I., & Povel, D.-J. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*(1), 61–69.

Silvia, P. J. (2005a). Cognitive appraisals and interest in visual art: Exploring an appraisal theory of aesthetic emotions. *Empirical Studies of the Arts*, *23*(2), 119–133.

Silvia, P. J. (2005b). Emotional responses to art: From collation and arousal to cognition and emotion. *Review of General Psychology*, *9*(4), 342–357.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science*, *17*(1), 57–60.

Silvia, P. J. (2010). Confusion and interest: The role of knowledge emotions in aesthetic experience. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(2), 75–80.

Silvia, P. J. (2012). Human emotions and aesthetic experience. In A. P. Shimamura, & S. E. Palmer (Eds.), *Aesthetic science: Connecting minds, brain and experience* (pp. 250–275). New York, NY: Oxford University Press.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Skorin-Kapov, J. (2015). *The aesthetics of desire and surprise: Phenomenology and speculation*. London: Lexington Books.

Skorin-Kapov, J. (2016). *The intertwining of aesthetics and ethics: Exceeding of expectations, ecstasy, sublimity*. London: Lexington Books.

Slater, A. (2007). "Escaping to the gallery": Understanding the motivations of visitors to galleries. *International Journal of Nonprofit and Voluntary Sector Marketing*, *12*(2), 149–162.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*(2), 110–120.

Sluckin, W., Hargreaves, D., & Colman, A. (1983). Novelty and human aesthetic preferences. In J. Archer, & L. I. A. Birke (Eds.), *Exploration in animals and humans* (pp. 245–269). Wokingham, England: Van Nostrand Reinhold.

Sohn, S., Seegebarth, B., & Moritz, M. (2017). The impact of perceived visual complexity of mobile online shops on user's satisfaction. *Psychology & Marketing*, *34*(2), 195–214.

Soley, G., & Spelke, E. S. (2016). Shared cultural knowledge: Effects of music on young children's social preferences. *Cognition*, *148*, 106–116.

Spee, B., Ishizu, T., Leder, H., Mikuni, J., Kawabata, H., & Pelowski, M. (2018). Neuropsychopharmacological aesthetics: A theoretical consideration of pharmacological approaches to causative brain study in aesthetics and art. In J. F. Christensen, & A. Gomila (Eds.), *Progress in brain research* (pp. 343–372). Amsterdam: Elsevier.

Spehar, B., Clifford, C. W., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics*, *27*(5), 813–820.

Spehar, B., & Taylor, R. P. (2013). Fractals in art and nature: Why do we like them? In B. E, Rogowitz, T. N. Pappas, & H. DeRidder (Eds.), *Human vision and electronic imaging XVIII*. Proceedings of SPIE, Volume 8651. Burlingame, CA: SPIE.

Spehar, B., Walker, N., & Taylor, R. P. (2016). Taxonomy of individual variations in aesthetic responses to fractal patterns. *Frontiers in Human Neuroscience*, *10*, 350.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393.

Stich, C., Eisermann, J., Knäuper, B., & Leder, H. (2007). Aesthetic properties of everyday objects. *Perceptual and Motor Skills*, *104*(3 suppl.), 1139–1168.

Street, N., Forsythe, A. M., Reilly, R., Taylor, R., & Helmy, M. S. (2016). A complex story: Universal preference vs. individual differences shaping aesthetic response to fractals patterns. *Frontiers in Human Neuroscience*, *10*, 213.

Streich, S. (2007). *Music complexity: A multi-faceted description of audio content*. Unpublished doctoral dissertation, University of Pompeu Fabra.

Sumpf, M., Jentschke, S., & Koelsch, S. (2015). Effects of aesthetic chills on a cardiac signature of emotionality. *PLoS ONE*, 10(6), 16.

Sun, L., Yamasaki, T., & Aizawa, K. (2018). Photo aesthetic quality estimation using visual complexity features. *Multimedia Tools and Applications*, *77*(5), 5189–5213.

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 370–381.

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7*(1), 14396.

Taylor, R. P., Spehar, B., Clifford, C. W. G., & Newell, B. R. (2008). The visual complexity of Pollock's dripped fractals. In A. A. Minai, & Y. Bar-Yam (Eds.), *Unifying themes in complex systems IV* (pp. 175–182). Berlin, Heidelberg, Germany: Springer.

Taylor, R. P., Spehar, B., Wise, J. A., Clifford, C., Newell, B. R., Hagerhall, C. M., . . . Martin, T. P. (2005). Perceptual and physiological responses to the visual complexity of fractal patterns. *Nonlinear Dynamics,Psychology, and Life Sciences*, *9*, 89–114.

Theiler, J. (1990). Estimating fractal dimension. *The Journal of the Optical Society of America A*, *7*(6), 1055–1073.

Tiihonen, M., Brattico, E., Maksimainen, J., Wikgren, J., & Saarikallio, S. (2017). Constituents of music and visual-art related pleasure—A critical integrative literature review. *Frontiers in Psychology*, *8*, 12.

Tinio, P. P. L., & Leder, H. (2009). Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica*, *130*(3), 241–250.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, *11*(4–5), 375–424.

Toussaint, G. T., & Beltran, J. F. (2013). Subsymmetries predict auditory and visual pattern complexity. *Perception*, *42*(10), 1095–1100.

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96–103.

Tseng, W. S.-W. (2018). Can visual ambiguity facilitate design ideation? *International Journal of Technology and Design Education*, *28*(2), 523–551.

Tuch, A. N., Bargas-Avila, J. A., Opwis, K., & Wilhelm, F. H. (2009). Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies*, *67*(9), 703–715.

Tuch, A. N., Presslaber, E. E., Stöcklin, M., Opwis, K., & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, *70*(11), 794–811.

Valenti, O., Mikus, N., & Klausberger, T. (2018). The cognitive nuances of surprising events: Exposure to unexpected stimuli elicits firing variations in neurons of the dorsal CA1 hippocampus. *Brain Structure & Function*, *223*(7), 3183–3211.

Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *i-Perception*, *2*(9), 1035–1062.

Van Den Bosch, I., Salimpoor, V., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience*, *7*, 534.

Van Geert, E., & Wagemans, J. (2020). Order, complexity, and aesthetic appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *14*, 135–154.

Vartanian, O. (2014). Empirical aesthetics: Hindsight and foresight. In P. P. L. Tinio, & J. K. Smith (Eds.), *The Cambridge handbook of the psychology of aesthetics and the arts* (pp 6–34). Cambridge: Cambridge University Press.

Vartanian, O., & Skov, M. (2014). Neural correlates of viewing paintings: Evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain and Cognition*, *87*, 52–56.

Venkatraman, A., Edlow, B. L., & Immordino-Yang, M. H. (2017). The brainstem in emotion: A review. *Frontiers in Neuroanatomy*, *11*, 12.

Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, *179*, 121–131.

Vessel, E. A., Starr, G. G., & Rubin, N. (2012). The brain on art: Intense aesthetic experience activates the default mode network. *Frontiers in Human Neuroscience*, *6*, 1–17.

Viengkham, C., & Spehar, B. (2018). Preference for fractal-scaling properties across synthetic noise images and artworks. *Frontiers in Psychology*, *9*, 1439.

Vinke, L. N. (2010). *Factors affecting the perceived rhythmic complexity of auditory rhythms* Unpublished Master's thesis, Bowling Green State University.

Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology*, *71*(1), 74–79.

Vladimirskiy, B., Urbanczik, R., & Senn, W. (2015). Hierarchical novelty-familiarity representation in the visual system by modular predictive coding. *PLoS ONE*, *10*(12), e0144636.

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig, Germany: Leopold Voss.

Vuust, P., Dietz, M. J., Witek, M., & Kringelbach, M. L. (2018). Now you hear it: A predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*, *1423*(1), 19–29.

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111.

Walker, E. L. (1980). Berlyne's theoretical contributions to psychology. *Motivation and Emotion*, *4*(2), 105–111.

Wang, M., & Li, X. (2017). Effects of the aesthetic design of icons on app downloads: Evidence from an android market. *Electronic Commerce Research*, *17*(1), 83–102.

Wang, Z., Duff, B. R. L., & Clayton, R. B. (2018). Establishing a factor model for aesthetic preference for visual complexity of brand logo. *Journal of Current Issues & Research in Advertising*, *39*(1), 83–100.

Wassiliwizky, E., Jacobsen, T., Heinrich, J., Schneiderbauer, M., & Menninghaus, W. (2017a). Tears falling on goosebumps: Co-occurrence of emotional lacrimation and emotional piloerection indicates a psychophysiological climax in emotional arousal. *Frontiers in Psychology*, *8*, 41.

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017b). The emotional power of poetry: Neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, *12*(8), 1229–1240.

Wilkins, R. W., Hodges, D. A., Laurienti, P. J., Steen, M., & Burdette, J. H. (2014). Network science and the effects of music preference on functional brain connectivity: From Beethoven to Eminem. *Scientific Reports*, *4*, 6130.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*, 483–494.

Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion*, *21*(1), 3–25.

Wundt, W. (1896). *Grundriss der Psychologie*. Leipzig, Germany: Wilhelm Engelmann.

Yeun, J. J., Eunsoo, B., & Jung, C. H. (2018). Managing the visual environment of a fashion store: Effects of visual complexity and order on sensation-seeking consumers. *International Journal of Retail & Distribution Management*, *46*(2), 210–226.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt. 2), 1–27.

Zatorre, R. J. (2015). Musical pleasure and reward: Mechanisms and dysfunction. *Annals of the New Academy of Sciences*, *1337*, 202–211.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558.

Zeki, S. (2002). Neural concept formation & art. Dante, Michelangelo, Wagner something, and indeed the ultimate thing, must be left over for the mind to do. *Journal of Consciousness Studies*, *9*(3), 53–76.

3. Discussion

3.1 Summary of main results

To effectively summarize the research findings with respect to the research questions, the research questions as phrased in the Introduction are restated below, followed by the insights we have gained from this doctoral thesis (see also Table 2):

- How does the distribution of images/excerpts in the two-dimensional emotion space spanned by arousal and pleasantness differ across environmental scenes, representational paintings and music?
  - For environmental scenes and representational paintings, the images followed a quadratic relationship (U-shaped distribution). Musical stimuli were more equally distributed and covered a circular or elliptical shape (Marin & Leder, 2013, Experiments 1-4).
  - For shorter presentation durations, these findings were replicated in Marin and Leder (2016) and Marin et al. (2016). Marin and Leder (2016) also showed in Experiments 1 and 2 that the shape of the affective spaces for presentation durations of 5 s and 25 s were more similar to each other than to the 1-s space. This was found to be true for both environmental scenes and representational paintings.

- How does the frequently reported positive linear relationship between complexity and arousal (controlled for familiarity) vary between the perception of affective environmental scenes, representational paintings and music?
  - The strength of the correlation varied considerably across stimulus sets for a presentation duration of 25 s, with the lowest positive correlation for environmental scenes (r_s = .27), followed by the one for representational paintings (r_s = .59), and the highest for solo piano music (r_s = .82) and solo piano music mixed with piano trios (r_s = .84) (Marin & Leder, 2013, Experiments 1-4).
  - Presentation duration affected this relationship for environmental scenes in females: no relationship for the 1-s condition, but significant positive relationships for the 5-s (r_s = .27) and 25-s conditions (r_s = .21) (Marin & Leder, 2016, Experiment 1).
  - Presentation duration affected this relationship for representational paintings in females, with stronger correlations for longer presentation durations: 1-s condition (r = .37), 5-s condition (r = .52) and 25-s condition (r = .49) (Marin & Leder, 2016, Experiment 2).

- The complexity-arousal relationship was largely unaffected by the nature of the hedonic tone ratings (pleasantness vs. liking vs. beauty) for a presentation duration of 5 s in a large group of females. For environmental scenes, the correlation was around r_s = .30, for cartoon-like environmental scenes around r_s = .28, for representational paintings around r_s = .47, and for solo piano music excerpts presented for 15 s around r_s = .86 (Marin et al., 2016, Experiments 1-2).

- Can the inverted U-shaped curve between complexity and pleasantness be observed across the three affective stimulus sets?
  - No, if familiarity is controlled for, the inverted U-shaped curve cannot be observed for the affective stimulus sets under investigation. There was no relationship between complexity and pleasantness for environmental scenes and representational paintings presented for 25 s. For solo piano music, complexity correlated positively (r_s = .42) among males but not in females. For solo piano music mixed with piano trio excerpts, there was a positive relationship among males (r_s = .41) and a non-significant negative relationship among females (Marin & Leder, 2013, Experiments 1-4).
  - Except for one significant positive correlation between ratings of complexity and pleasantness of environmental scenes presented for 5 s, there were no significant relationships observed for environmental scenes presented for 1 and 25 s and representational paintings presented for 1, 5 and 25 s (Marin & Leder, 2016, Experiments 1-2).
  - The relationship between complexity and pleasantness did not follow an inverted U-shaped curve when environmental scenes, environmental scenes converted to cartoons, and representational paintings were presented for 5 s, and when solo piano music was presented for 15 s to a group of females. However, an inverted U-shaped curve was observed for complexity and beauty ratings of solo piano music (Marin et al., 2016, Experiments 1-2).

- Does subjective arousal act as a mediator in the complexity-hedonic tone relationship in all three affective stimulus sets when familiarity is controlled for?
  - Yes, subjective arousal is a mediator in the complexity-hedonic tone relationship. In males and females, arousal was identified as a significant mediator in the complexity-pleasantness relationship for environmental scenes. At all levels of picture complexity, pleasantness decreased when complexity increased and affected arousal. For representational paintings, significant indirect effects

of arousal on pleasantness only emerged either when familiarity was not considered in the model, or when both genders were analyzed separately (only for pictures of moderate and high complexity). There were no direct effects of complexity on pleasantness for either type of visual stimuli. For solo piano music, arousal was a significant mediator, and the effect observed for both genders was mostly driven by females. Pleasantness decreased linearly as complexity increased. There was also a significant direct effect of complexity on pleasantness. These results were replicated for solo piano music mixed with piano trios (Marin & Leder, 2013, Experiments 1-4).

- For environmental scenes, the results of Marin and Leder (2013) were replicated among female participants for presentation durations of 5 and 25 s (but not for the 1-s condition). An increase in complexity led to a decrease in pleasantness through a significant effect of complexity on arousal. For representational paintings of moderate and high complexity, there was a significant indirect effect of complexity on pleasantness through arousal for presentation durations of 1 s, 5 s and 25 s. There were no mediation effects for paintings of low complexity (Marin & Leder, 2016, Experiments 1-2).

- Do concepts such as pleasantness, liking and beauty, subsumed under the term *hedonic tone*, relate similarly to complexity and arousal across the three affective stimulus sets?
  - No, the correlation patterns differed across the stimulus sets. In environmental scenes and environmental scenes converted into cartoons, pleasantness, liking and beauty did not correlate with complexity. In representational paintings, however, these three concepts correlated differently with complexity when familiarity was controlled for: positive correlation between beauty and complexity, no correlations between complexity and liking and pleasantness, respectively (Marin et al., 2016, Experiment 1).
  - In environmental scenes and environmental scenes converted into cartoons, complexity and arousal correlated in a similar fashion with each other when familiarity was controlled for (r_s ~ .30). In representational paintings, this relationship was somewhat stronger (r_s ~ .46) across the three hedonic tone measures (Marin et al., 2016, Experiment 1).
  - For solo piano music, complexity correlated negatively with pleasantness and positively with liking when familiarity was controlled for. For beauty ratings, we observed a significant inverted U-shaped curve with complexity. Arousal was

highly correlated with complexity (r_s ~ .86) across all three hedonic tone conditions (Marin et al., 2016, Experiment 2).

- How does the performance of objective measures of visual complexity depend on the stimulus type (environmental scenes vs. representational paintings)?
  - The performance of objective measures of visual complexity depends on the stimulus type. For a set of 13 measures of visual complexity, compression algorithms and RMS contrast measures correlated well with subjective complexity ratings of environmental scenes (moderate effect size). Compressed file size did not correlate in a similar way with subjective complexity in representational paintings, but phase congruency was a significant predictor. In both stimulus sets, edge detection algorithms predicted visual complexity better than compressed file size (Marin & Leder, 2013, Experiments 1-2; 2016, Experiments 1-2; 2022, Experiment 1).
  - In representational paintings, the Canny algorithm was sensitive to brushstrokes in the background of paintings and thus correlated negatively with subjective complexity (Marin & Leder, 2013, Experiment 2; Marin & Leder, 2016, Experiment 2).
  - In representational paintings, the RMS contrast measures also correlated positively with subjective arousal in the 5-s condition. Correlations of objective measures of complexity with arousal were not observed for environmental scenes (Marin & Leder, 2016, Experiments 1-2).
  - Objective measures of complexity are sensitive to subdimensions of visual complexity. In environmental scenes presented for 1 s, correlations between the number of objects and all measures of objective complexity (except for the phase congruency measures) were stronger than those between general complexity ratings and the respective objective measures. The same result was obtained for figure-ground vs. complex scene ratings. Disorganization ratings correlated less strongly with the objective measures than ratings of the number of objects and figure-ground vs. complex scene. In the 5-s condition, objective measures of complexity correlated similarly well with ratings of general complexity, the number of objects and figure-ground vs. complex scene. The correlation strength between disorganization ratings and the objective measures were unaffected by presentation duration (Marin & Leder, 2016, Experiment 3).

- In representational paintings, the multidimensionality of complexity became less obvious through correlations with objective measures. In the 1-s condition, the perimeter and Canny edge detection algorithms and entropy were not sensitive to subdimension of complexity, but some of the RMS contrast and phase congruency measures showed differential effects. In the 5-s condition, the Canny algorithm and phase congruency measures correlated less with disorganization than with the other type of complexity ratings. In general, the objective measures showed weaker correlations with disorganization in the 5-s condition than in the 1-s condition (Marin & Leder, 2016, Experiment 3).
- The pattern of intercorrelations of nine objective complexity measures (including spatial frequency) differed between environmental scenes and representational paintings (Marin & Leder, 2022, Tables S19-S20).
- Different types of objective measures predicted average fixation counts for environmental scenes and representational paintings (Marin & Leder, 2022, Experiment 1).

- Can objective measures of visual complexity (based on Kolmogorov complexity and edge detection) in their analogous form be applied to the study of musical complexity?
  - Yes, compression file size (MP3, Ogg Vorbis and FLAT) and edge detection were found to be significant predictors of musical complexity of 25-s excerpts of two different musical styles, with an r_s ranging between .35 and .58 for solo piano music and an r_s between .37 and .79 for solo piano music mixed with piano trios (Marin & Leder, 2013, Experiments 3-4).
  - Objective measures of musical complexity also correlated significantly with subjective arousal (Marin & Leder, 2013, Experiments 3-4).

- How does presentation duration (1 s, 5 s and 25 s) affect the relationship between objective and subjective complexity in the visual domain?
  - For environmental scenes, objective measures of complexity predicted subjective complexity similarly well for presentation durations of 1, 5 and 25 s (although there was a non-significant increase of correlation strength with longer presentation duration) (Marin & Leder, 2016, Experiment 1).
  - For representational paintings, several objective measures correlated (non-significantly) more strongly in the 1-s condition than in the longer conditions, whereas for other measures the correlations with subjective complexity were

similar across the three conditions of presentation duration (Marin & Leder, 2016, Experiment 2).

- When subdimensions of visual complexity were correlated with objective complexity measures in environmental scenes and representational paintings, these correlations were generally unaffected by presentation durations involving 1 and 5 s (Marin & Leder, 2016, Experiment 3).

- When controlling for familiarity, how does presentation duration (1 s, 5 s and 25 s) affect the relationships between complexity, pleasantness and arousal in the visual domain?
  - In environmental scenes, the complexity-arousal relationship was only significant for the 5- and 25-s conditions. The positive complexity-pleasantness relationship was only present in the 5-s condition (Marin & Leder, 2016, Experiment 1).
  - In representational paintings, the relationship between complexity and arousal was lower for the 1-s condition than for the 5-s and 25-s conditions. The complexity-pleasantness relationship did not reach significance for any of the three conditions of presentation duration (Marin & Leder, 2016, Experiment 2).

- How does presentation duration (1 s, 5 s and 25 s) affect mean ratings of complexity, pleasantness and arousal in the visual domain?
  - In environmental scenes, familiarity increased over time, and arousal ratings were higher in the 5- and 25-s conditions than in the 1-s condition. Pleasantness ratings were the highest in the 5-s condition and lower in the 1- and 25-s conditions. Complexity ratings also followed an inverted U-shaped relationship across the three presentation durations, and the lowest ratings were observed for the 25-s condition (Marin & Leder, 2016, Experiment 1).
  - In representational paintings, familiarity ratings were higher in the 1- and 25-s conditions than in the 5-s condition. Complexity ratings were the highest in the 5-s condition with lower ratings for the 1-s and 25-s conditions. Arousal ratings were higher in the 1-s and 25-s conditions than in the 5-s condition. Pleasantness ratings were the highest in the 5-s conditions and lower in the two other conditions (Marin & Leder, 2016, Experiment 2).

- When controlling for familiarity, how do personality traits related to affect and art/music interest explain individual differences regarding the processing of

complexity and arousal when perceiving environmental scenes, paintings and music?

- Complexity ratings of environmental scenes were unrelated to affective personality traits. On the other hand, complexity ratings of environmental scenes converted into cartoons were positively related to stress reactivity, and complexity ratings of representational paintings were related to stress reactivity and empathy (Marin & Leder, 2018, Study 1).
- Affective personality traits predicted arousal ratings to a larger degree than complexity ratings. Arousal ratings of environmental scenes were predicted by stress reactivity, empathy and trait emotional intelligence. Arousal ratings of cartoons and representational paintings were predicted by trait emotional intelligence and empathy (Marin & Leder, 2018, Study 1).
- Art interest did not predict visual complexity and arousal in the visual domain (Marin & Leder, 2018, Study 1).
- Musical complexity was predicted by stress reactivity, trait emotional intelligence, emotional self-efficacy and musical sophistication. For arousal ratings, the same personality traits were predictive, but not musical sophistication (Marin & Leder, 2018, Study 2).

- How does gender affect the relationships between familiarity, complexity, arousal and pleasantness, respectively?
  - Gender affected the inter-relationships between familiarity, complexity, arousal and pleasantness. Gender effects were stronger for artistic stimuli than for environmental scenes and gender affected the direction (not only the strength) of associations between these variables (Marin & Leder, 2013, Experiments 1-4).

- How does viewing behaviour differ between the perception of affective environmental scenes and representational paintings?
  - For a presentation duration of 25 s, representational paintings received a higher number of fixations counts (and shorter mean fixation durations) than environmental scenes. There was also a significant pleasantness × picture category interaction for saccade amplitudes, with longer saccade amplitudes for more pleasant paintings. For a presentation duration of 6 s, picture category interacted with pleasantness when predicting fixation counts and fixation durations. Pleasant paintings received a lower number of fixation counts (and shorter mean fixation durations) than environmental scenes (Marin & Leder, 2022, Experiments 1-2).

- Pleasantness, but not visual complexity, leads to opposite effects on eye movement parameters when comparing environmental scenes with representational paintings: fixation durations were longer for pleasant paintings and shorter for pleasant environmental scenes (Marin & Leder, 2022, Experiments 1-2).

- How do arousal, pleasantness and complexity affect viewing behaviour? Is there an effect of presentation duration (long vs. short)?
  - Arousal only positively predicted eye movement parameters when viewing environmental scenes for 25 s. Arousal had no effect on eye movement parameters when viewing representational paintings (Marin & Leder, 2022, Experiments 1-2).
  - For a presentation duration of 25 s, more complex and more pleasant environmental scenes were associated with a larger number of fixations (and shorter fixation durations). An increase in pleasantness led to shorter saccade amplitudes. These findings were replicated for a presentation duration of 6 s (Marin & Leder, 2022, Experiments 1-2).
  - For a presentation duration of 25 s, a smaller number of fixations (and longer fixation durations) were detected when viewing more pleasant representational paintings. Saccade amplitudes were unaffected by complexity and pleasantness. The effects were similar for the 6-s condition, but with an additional effect of complexity on fixation counts and fixation durations (Marin & Leder, 2022, Experiments 1-2).
  - Picture category only emerged as a significant fixed effect in the 25-s condition.

Table 2

Overview of the Main Findings of this Doctoral Thesis

Feature	Environmental scenes	Representational paintings	Music
Arousal-pleasantness emotion space	Quadratic*	Quadratic*	Circular/elliptical shape
Complexity-arousal relationship	Weak*	Strong*	Strong
Arousal as a mediator in the complexity-pleasantness relationship	Yes, and no direct effect of complexity on pleasantness*	Yes, and no direct effect of complexity on pleasantness	Yes, and also direct effect of complexity on pleasantness
Inverted U-shaped curve between complexity and pleasantness	No relationship*	No relationship	Positive relationship for males
Kolmogorov measures of complexity	Yes	No	Yes
Edge detection measures of complexity	Yes	Yes	Yes
Sensitivity of objective measures to subdimensions of complexity	Yes	Yes	NA
Effects of presentation duration on the interrelationships between familiarity, complexity, pleasantness and arousal	Yes*	Yes*	NA
Change of mean complexity, arousal and pleasantness ratings across presentation durations of 1, 5 and 25 s	Yes	Yes	NA
Multidimensionality of hedonic tone, different correlation patterns with complexity	No	+ Beauty	Inverted U-shaped curve for beauty + Liking - Pleasantness
Affective personality traits and their effect on subjective complexity	No	Yes	Yes
Affective personality traits and their effect on subjective arousal	Yes	Yes	Yes
Effects of art interest / musical sophistication on subjective complexity and/or arousal	NA	No	Yes
Gender effects on the interrelationships between familiarity, complexity, pleasantness and arousal	Yes	Yes	Yes
Effect of complexity on eye movement parameters	Yes	Yes*	NA
Effect of arousal on eye movement parameters	Yes*	No	NA
Effect of pleasantness on eye movement parameters	Yes	Yes	NA

Note. *Additional significant effects of presentation duration.

3.2 Berlyne revisited: Arousal and complexity still rule

Arousal and complexity are key parameters in Berlyne's psychobiological model of aesthetic response (1971). Recent theoretical endeavours in empirical aesthetics have frequently ignored the potential role of arousal during aesthetic episodes, although researchers have shown unceasing interest in the complexity-hedonic tone relationship of the arousal-boost mechanism described by Berlyne (for a review see Marin, 2022). By design we selected affective stimuli that also included unpleasant emotions (Silvia, 2009), thus following a comprehensive approach that acknowledges that aesthetic judgements depend both on aesthetic emotions and other aesthetic qualities (Stamkou et al., 2024). The current data provided convincing evidence for a positive relationship between subjective complexity and arousal in responses to visual and musical stimuli sets when controlling for familiarity (another collative variable in Berlyne's framework), which is in line with Berlyne's theory (Marin & Leder, 2013, 2016; Marin et al., 2016). We demonstrated that the strength of this relationship seems to depend on the stimulus type, with the weakest relationship observed for environmental scenes and the strongest for music, which is relevant for theories of aesthetic experiences. Furthermore, Berlyne predicted that arousal acts as a mediator in the complexity – hedonic tone relationship. We also found support for this hypothesis in our studies when ratings of environmental scenes, representational paintings and music were considered (Marin & Leder, 2013). However, when investigating the impact of arousal on eye movements during the free exploration of environmental scenes and representational paintings, we only observed behavioural arousal effects for environmental scenes but not for representational paintings (Marin & Leder, 2022). Visual complexity predicted eye movement parameters well across stimulus types and presentation durations, which again demonstrates the well-documented relevance of structural variables in the perception of semantic scenes (Sun & Firestone, 2021; Wu et al., 2014). Subjective arousal also correlated negatively with familiarity across stimulus sets, which is in line with Berlyne's theory: novelty (unfamiliarity) is associated with heightened arousal (Berlyne, 1970, but see also Van Den Bosch et al., 2013, for contrary results for music). In short, at least two core components of Berlyne's theory centred around the concept of arousal could be supported by our empirical findings based on subjective ratings: the positive relationship between complexity and arousal and arousal as a mediator in the complexity-hedonic tone relationship.

Following Berlyne's approach, it was also possible to demonstrate (with state-of-the-art image and sound analyses) that "formal" stimulus properties such as complexity can be predicted by objective measures, which generally correlated well with subjective complexity. Arousal in response to representational paintings and music also correlated with some of our objective complexity measures. Nevertheless, our results yielded no signs of the predicted inverted-U-shaped relationship between complexity and pleasantness by Berlyne (1971). Considering different measures of hedonic tone (liking, beauty and pleasantness), we demonstrated that this type of relationship only occurs for complexity and beauty ratings of solo piano music (Marin et al., 2016).

Our current findings thus call for a differentiated view on Berlyne's psychobiological model (1971). It seems that the link between arousal and collative variables, such as complexity and familiarity, is a robust finding present in the visual and auditory domains. Therefore, one can argue that this part of Berlyne's theory should not be abandoned and internal affective state scales should receive more attention in theoretical models. If this finding is replicated, for example by studying large non-affective stimuli sets (including abstract visual stimuli) and different participant groups, the correlation between complexity and arousal could be integrated into other models of aesthetic experience or aesthetic response. Complexity is already one of several stimulus-related variables in various models, but its specific link to arousal (and familiarity) has not been of great relevance outside Berlyne's psychobiological model, although several models incorporate affective responses more broadly: Pelowski, Markey, et al. (2017) integrated arousal and affective responses into their integrated model of top-down and bottom-up processes in art perception by discussing affect with regard to different processing output scenarios. Osgood et al. (1957) already regarded Activity and Potency as central dimensions in the generation of meaning (semantic differential). Arousal and complexity could be easily integrated into the network model of aesthetic effects proposed by Specker et al. (2021), which is based on the semantic differential approach. Lively vs. still is another aesthetic effect that is somewhat related to arousal and activity, but complexity is not explicitly mentioned by Specker et al. (2021). Another example is the pleasure-interest model of aesthetic liking proposed by Graf and Landwehr (2015), in which complexity and arousal may be integrated relatively easily into the existing framework based on fluency theory. Similarly, theories of aesthetic experiences that are grounded in predic-

tive coding (Frascaroli et al., 2024; Van de Cruys et al., 2024; Van de Cruys & Wagemans, 2011) may also profit from integrating arousal responses to stimulus complexity. Mandler (2003) links arousal to prediction errors, and more recently, researchers have suggested that the brain makes predictions about stimulus features as well as about the affective contents of stimuli, the latter being related to arousal responses in the case of prediction errors (Ferreira-Santos, 2016). In short, most current theories on aesthetic experiences somehow link cognitive and affective processing (see also Miller & Hübner, 2020), which are thought to underlie general perception and information processing (Barrett & Bar, 2009), but most of these theories do not explicitly mention arousal as the affective dimension underlying pleasure or displeasure (but see Blood & Zatorre, 2001, who reported a link between arousal and musical pleasure). However, a recent study on the appreciation of Chinese opera (Li et al., 2024) has integrated psychophysiological arousal into a wider framework involving other core variables of aesthetic experiences (familiarity and novelty). Such studies reflect Berlyne's legacy and may motivate to acknowledge the role of arousal to a greater extent as it is currently the case. There is a current interest in studying why humans are curious, seek knowledge and enjoy exploring new information. For example, Cotter et al. (2022) examined motivations for visiting an art museum. More generally, it may be useful to the field to reflect on whether a motivational approach to aesthetics, especially in the context of evolution, may be a fruitful path for the future (see also Menninghaus et al., 2019).

Another line of research considers arousal when studying aesthetic emotions. This line of research takes a narrower perspective and mostly focuses on affective responses without necessarily considering cognitive contributions to aesthetic experiences. For instance, researchers have reported on aesthetic chills and arousal responses (Jain et al., 2022), arousal as a dimension in aesthetic emotions (Beermann et al., 2021) and as an important dimension in aesthetic experiences of live concerts (Merrill et al., 2021). Other studies have been interested in the emotional contents of stimuli (core affect), mostly in applied contexts (Bhandari et al., 2019; Gregorians et al., 2022). Clearly, these studies have shown that arousal and more complex emotions are associated with aesthetic experiences (and several models acknowledge this, see e.g., Brattico et al., 2013; Leder et al., 2004; Marković, 2012; Pelowski, Markey, et al., 2017; Silvia, 2005; Specker et al., 2021). Future studies will have to isolate the effects of arousal induced by the affective contents of

stimuli (e.g., induced by semantic contents in the visual domain) from those arousal effects that can be linked to structural features of the stimuli, such as complexity (Fan et al., 2020; Redies et al., 2020). This may be possible in the visual domain by using controlled stimuli as well as psychophysiological and subjective measures of arousal. In the auditory domain, such a differentiation may be possible for vocal music but not for instrumental music. Furthermore, it is a well-established fact that affect can influence all types of cognitive processing (e.g., Storbeck & Clore, 2008); therefore, the affect-cognition relation during aesthetic processing deserves more attention, although it is difficult to study and model the exact underlying neuronal and psychological mechanisms.

As elaborated in depth in Marin (2022), support for the predicted inverted U-shaped relationship by Berlyne (1971) (e.g., Chmiel & Schubert, 2019; Delplanque et al., 2019) as well as counter-evidence are currently reported (e.g., Althuizen, 2021). There may be several underlying reasons for these discrepant findings (Marin, 2022). Some of them may be grounded in the nature of the stimuli and how they were selected. Importantly, our current data based on affective stimuli suggests that it is necessary to consider the multidimensionality of complexity (Nadal et al., 2010) alongside the multidimensionality of hedonic tone (Marin et al., 2016) to fully understand the nature of the complexity-hedonic tone relationship. Therefore, concepts such as liking, beauty and pleasantness should not be subsumed under one common term as suggested by Berlyne (1971). The multidimensionality of hedonic tone becomes also apparent in models of aesthetic experiences, such as in Brattico et al. (2013), in which concepts such as liking, beauty and pleasantness have their own role and underlying neurophysiology when predicting musical aesthetic experiences. In the visual domain, current models also differentiate between these concepts (Pelowski, Markey, et al., 2017). The differentiation between liking, beauty and pleasantness when perceiving works of art is also relevant in the context of theories of motivation and reward (see Marin, 2022).

The field has recently seen a growing interest in the study of individual differences, for instance, pointing out the role of personality, expertise and aesthetic sensitivity (Augustin & Leder, 2006; Celikors & Field, 2025; Kozbelt, 2022; Meyer et al., 2023; Myszkowski, 2022; Swami & Furnham, 2022; Tetzlaff et al., 2024). Berlyne (1974) already discussed the important role of expertise in aesthetic responses. Recent studies have shown that individual differences in preferences can explain dis-

crepancies regarding his inverted U-shaped relationship for visual and musical stimuli (Güçlütürk et al., 2016; Güçlütürk & van Lier, 2019) and that preferences for musical styles can predict preferences for motives in visual art (Ercegovac et al., 2015). Other research has compared how the Big five personality traits are related to preferences for art forms, such as music, architecture and visual art (Cleridou & Furnham, 2014; Ercegovac et al., 2015), and found that openness is very predictive across art forms. More recently, Miller and Hübner (2023) reported gender effects regarding empathy, aesthetic judgements and aesthetic inferences in response to visual art.

In this doctoral thesis, we studied how affective personality traits and art interest/musical sophistication are related to ratings of arousal and complexity (Marin & Leder, 2018). Moreover, we considered gender and familiarity in our analyses and research designs (Marin & Leder, 2013, 2016, 2018, 2022; Marin et al., 2016). As has been noted, we found several effects of individual differences, regarding single ratings and associations between variables, and we showed that, as predicted, affect-related personality traits have a larger impact on arousal than complexity ratings for affective visual stimuli, and that this difference was not present for music. Art interest was not predictive of visual complexity and arousal (van Paasschen et al., 2015, but see Tomasz Jankowski et al., 2020), but increasing musical sophistication was associated with higher complexity ratings (but see Orr & Ohlsson, 2005). This has far-reaching implications for theory-building because it may be that specific person-related factors show differential effects on various core variables of aesthetic experience in visual compared to auditory or multi-modal art forms. Furthermore, the degree of individual differences on aesthetic responses may also depend on the processing stage if one considers models such as those by Leder et al. (2004) and Pelowski, Markey et al. (2017), and one can surmise that the effects may be larger at later stages than at early sensory processing stages. Appraisal theorists have argued that appraisal mechanisms can explain individual differences in affective and aesthetic responses better than other theories (Silvia, 2006). Only systematic investigations considering different stimulus types will eventually offer a clear picture. Moreover, the field needs to reflect on how expertise is conceptualized and measured (Sorek et al., 2024). At the moment, we are only starting to get a better understanding of individual differences in aesthetic experiences, especially due to the large number of possible facets (in terms of individual differences and stimulus material) to be studied.

Another timely topic is studying the time course and dynamics of aesthetic experiences (e.g., Belfi et al., 2019; Palczak et al., 2020; Spee et al., 2024; Wagner et al., 2021). We employed different presentation durations (between 1 and 25 s) when presenting visual stimuli to different participant groups and found, in the context of Berlyne's theory, that mediation effects of arousal are only present for longer presentation durations when viewing environmental scenes, that correlations between objective measures of complexity and subjective complexity ratings are partly dependent on presentation duration for representational paintings, and that correlations between objective measures of complexity and subdimensions of visual complexity are more stable across presentation durations than those with general complexity ratings (Marin & Leder, 2016). Moreover, the complexity-arousal relationship was not present for the 1-s condition when viewing environmental scenes and was weaker for the 1-s condition than in the longer conditions when viewing representational paintings. In general, inter-relationships between familiarity, arousal, complexity and pleasantness were stronger for longer presentation times. We also noted interesting effects of presentation duration (6 vs. 25 s) when examining eye movement parameters (Marin & Leder, 2022): picture category only emerged as a predictor for the 25-s condition. Thus, the current data provide convincing evidence that time is a key factor when studying visual aesthetic experiences, and that it is worth studying longer exposure times besides initial perceptual processing stages in the future, especially those around 17 s, for which Brieber et al. (2020) reported the highest appreciation scores. In the long run, we will need to investigate temporal processing in the visual domain (both for static and dynamic stimuli) in greater detail, not just in the field of neuroaesthetics but also in general psychology, which will ultimately foster the development of theories that acknowledge temporal aspects to a larger extent than previous theories did (see e.g., Husselman et al., 2024; Vara Sánchez, 2022).

Our current results regarding Berlyne's theory are not generalizable beyond the mostly female samples. Thus, future research will have to investigate the role of gender in relation to responses to affective artworks in greater depth (Marin & Leder, 2013), also in relation to individual differences (Miller & Hübner, 2023). We studied genuine artworks and our visual stimuli were only mildly controlled (no highly erotic stimuli etc.) to keep the ecological validity high. In his comprehensive approach, Berlyne (1971) proposed to study both highly controlled stimuli and actual artworks. Thus, one could extend our approach by differentiating between various semantic

categories (e.g., Löw et al., 2013; Zhou et al., 2023) as well as by studying abstract artworks, other musical styles and employing artificial, highly-controlled musical and visual stimuli. More importantly, future research should include psychophysiological measures of arousal, perhaps in combination with eye tracking, as well as other subjective ratings that are of relevance to Berlyne's theory and empirical aesthetics, such as interest, curiosity and separate ratings for perceived vs. felt emotions. It would be also worth studying other types of behavioural responses besides eye movements, such as viewing/listening time. Berlyne's *New Experimental Aesthetics* (1974) employed a very comprehensive research approach, which could be easily extended with the current stimulus sets. One could also question whether subjective arousal is a uniform concept and study tension arousal as well as energy arousal (Ilie & Thompson, 2006). For instance, Gingras et al. (2015) reported a high correlation between energy and tension arousal of Romantic piano trio excerpts, but this needs to be confirmed for other musical styles and art forms. Last, the neurophysiological basis of arousal is complex and multiple brain circuits comprising the brainstem, forebrain and cortex have been identified to contribute to subjective arousal (Berridge, 2008; Maness et al., 2022; Sara, 2009; Quadt et al., 2022, see also Marin, 2022), which will pose a challenge to the development of theories on aesthetic responses.

3.3 Emotion across sensory and artistic domains

The comparative approached followed in this doctoral thesis increased our understanding of emotion (core affect) and complexity in the visual and auditory domains. The self-selected stimuli were meant to differ in arousal, pleasantness and complexity (note that much effort was put into the stimulus selection procedure to accomplish this goal), and the subjective ratings have confirmed different types of two-dimensional emotion spaces for visual versus musical stimuli. For visual stimuli with semantic and affective contents (excluding highly erotic stimuli and those showing mutilation), the relation between arousal and pleasantness followed a U-shape (Codispoti et al., 2009; Lang et al., 1997), whereas for musical stimuli, the stimuli were more equally distributed in the arousal-pleasantness space (Campos-Bueno et al., 2015; Grekow, 2016; Schubert, 1999). These findings were replicated for different presentation durations and participant groups (Marin & Leder, 2013, 2016).

Whether and how arousal and pleasantness are correlated with each other in the visual domain is still a matter of debate (Colombetti & Kuppens, 2024; Haj-Ali et

al., 2020; Kuppens et al., 2013; Petrolini & Viola, 2020). The U-shape function frequently reported for environmental scenes (which we also observed for representational paintings) may suggest that arousal and pleasantness are not clearly separable (Dan-Glauser & Scherer, 2011; Kurdi et al., 2017; Lang et al., 1997). Similar evidence was provided by Kron et al. (2015) who found that subjective arousal was not a better predictor of electrodermal responses to visual stimuli than subjective valence. Our current findings somewhat support this view. Our mediated regression analyses (complexity-arousal-pleasantness) suggest that arousal and pleasantness are related with each other (which supports Berlyne's theory (1971) but not Russell's (1980), who regarded arousal and pleasantness as independent dimensions). This held true for all our studies involving either visual or musical stimuli, but one cannot conclude from our current data that arousal and pleasantness are similar or identical concepts. For instance, our current objective measures of complexity correlated differently with arousal and pleasantness, for representational paintings and music. Moreover, in our eye movements study (Marin & Leder, 2022), we observed independent effects of arousal and pleasantness. We did not obtain individual ratings of arousal and pleasantness by the participants whose eye movements were measured, but instead used ratings from other participants to predict eye movements. Future studies may collect subjective ratings by the same participants to further elucidate the relationship between arousal and pleasantness when viewing natural scenes (Kuppens et al., 2013) and paintings. More generally, future research will have to meet the challenge of including various types of visual and musical stimuli and employing a multi-method approach to examine how and whether arousal and pleasantness are related with each other at the psychological and physiological levels.

It is difficult to interpret differences regarding the distribution of the stimuli and the shape of the emotion spaces across the visual and auditory domains. On the one hand, arousal and pleasantness may be less related in music and there may be several reasons for this: as Eerola and Vuoskoski (2013) pointed out in their review on musical emotion models, music does generally not induce high-arousing negative emotions (see also Mehl et al., 2023), which was also observed in our studies using 19th-century piano music. Moreover, music does usually not communicate semantic contents, and it is possible that this is the main reason for the different shapes of the current emotion spaces. Future studies may include affective environmental sounds (e.g., IADS, Yang et al., 2018) and directly compare them to music to study the role

of semantic contents in auditory affective communication. Campos-Bueno et al. (2015) also compared the arousal-valence space of 52 paintings of different styles based on a presentation duration of 3 s with the one of 52 musical excerpts of an average length of 15 s. Their results were similar to our findings: the musical excerpts were more dispersed than the paintings, and the emotion space covered by the paintings was similar to the one observed for IAPS pictures. Altogether, this suggests that the arousal-valence spaces may differ across art forms and across sensory domains more generally.

Our findings offered also insights into how the distribution of visual stimuli in the emotion space is depending on the presentation duration: the emotion spaces of visual stimuli presented for 5 and 25 s were more similar to each other than the emotion spaces for the 1-s condition, suggesting that affective ratings may change over time (Marin & Leder, 2016). Since musical information expands over time, studies with musical stimuli of different length are not directly comparable to studies involving static visual stimuli of different presentation duration. However, if the musical contents is not changing much over time, it may be worthwhile running a study with musical excerpts of different lengths in the future. Studies have demonstrated that excerpts of around 100-250 ms length are sufficient to recognize musical emotions (Bigand, et al., 2005; Nordström & Laukka, 2019), and thus it may be informative to explore the dynamics of the emotion space of 19th century music. As music, visual perception is also a dynamic process, even when viewing static images (Hegdé, 2008), because eye movements change over time and thus determine the type and amount of visual information that is processed. Therefore, a comparison between musical and visual processing is justifiable on psychological grounds.

Recent studies have also shown interest in the question of how emotions can be conceptualized across artistic art forms, such as paintings and music. For instance, Miu et al. (2016) used a music-specific emotion scale to assess emotional responses to music and paintings. They did not present any real stimuli and participants gave ratings based on their memory. However, some interesting findings emerged that corresponded to previous findings. For example, they found that music evokes a different spectrum of emotions than paintings, with different frequencies for several categories. Differential results were also noted for the contributing factors to the emotional experience (i.e., stimulus features, previous knowledge, prior mood etc.). Interestingly, emotions induced by music were rated as being more similar to

everyday emotions than those induced by paintings. Van Der Lee and Van Enschot (2022) were also inspired by music research and examined whether the mechanisms that have been documented to induce musical emotions are also relevant for emotions induced by paintings. This is a timely approach because it will help elucidate the question of how aesthetic experiences are induced, of which emotions are an integral part of. The authors modified the original BRECVEMA mechanisms proposed by Juslin (2013), referring to brain stem reflex, rhythmic entrainment, evaluative conditioning, emotional contagion, visual imagery, episodic memory and musical expectancy, and developed the six SECMEA mechanisms. Rhythmic entrainment was excluded for the visual domain and instead a new mechanism based on appraisal was added. The authors summarized and discussed tentative empirical evidence for the respective mechanisms, which motivates further research along these lines. Tiihonen et al. (2017) reviewed how pleasure in music and visual art is conceptualized and also reported interesting findings regarding the role of emotions: it seems as if pleasure is rooted in core affect when listening to music and in cognitive processing when experiencing visual art. More recently, Mehl et al. (2023) studied in an online study aesthetic responses to music, paintings and poetry of different styles and time periods in a between-subjects design. They predicted aesthetic appeal by ratings of vividness of evoked imagery, valence (positive. vs. negative) and arousal. Across all stimulus sets, vividness was the strongest predictor. Valence was a strong predictor for music, and the authors observed a U-shaped relationship between arousal and aesthetic appeal. Bruns et al. (2023) examined how beauty judgements are determined by the affective contents of images and songs and the perceivers' mood states. They reported interesting differences between images and music, which were also influenced by presentation duration. Altogether, these studies based on subjective reports followed a comparative approach (see also Marin, 2015; 2022) to get insights into the nature of (aesthetic) emotions, mostly focusing on emotion categories, whereas in this doctoral thesis, core affect was the centre of interest (see also Mehl et al., 2023). These two approaches are complementary and welcome, especially because it is yet unclear what type of emotion model (categorical vs. dimensional etc.) is most appropriate in emotion science and how emotions are actually defined (Sander, 2013; Šimić et al., 2021).

Across our experiments, we noted that familiarity correlated moderately to strongly with arousal, complexity and particularly with pleasantness. Familiarity is not

only a key variable in Berlyne's framework (1971), but also gained importance among appraisal theorists in the context of emotion (Ellsworth & Scherer, 2003). When subjects interpret the personal significance of events, familiarity vs. novelty is an important dimension alongside the intrinsic pleasantness of objects or events. How emotions are elicited is also a matter of debate in the field, but our findings suggest that research on emotions in the context of empirical aesthetics should incorporate familiarity because it seems to correlate with both core affect and complexity. For instance, within the framework of Berlyne's theory, Yanagisawa et al. (2019) conducted an EEG-study and modelled prediction errors during listening to percussive music following a Bayesian approach. Instrumental familiarity was shown to affect arousal responses and information processing in the brain in line with Berlyne's theory.

Theory building depends not only on identifying the relevant variables to explain a phenomenon, but also on controlling for those that assert an influence on all other variables of interest and can explain individual differences. Familiarity appears to be such a variable that can relatively easily be assessed and statistically controlled for (see also Fekete et al., 2023). Familiarity and its correlation with core affect may also contribute to our understanding of whether concepts such as valence and arousal show domain-specific or domain-general characteristics. For instance, the correlation strength between familiarity and pleasantness was between $r = .6$ and $r = .8$ across our visual and auditory stimulus sets, but the one between familiarity and arousal was negative, much weaker and showed a larger variability. There may be several reasons for this observation. In an evolutionary context, one could surmise that since pleasantness (valence) is a dimension that largely determines approach-avoidance behaviour (Mauss & Robinson, 2009), valence may be generally significant to humans and similarly processed across sensory domains (Gu et al., 2019; Kim et al., 2017; Park & Kim, 2024), which may also be reflected by domain-general relations to familiarity/novelty. This hypothesis could be tested in greater depth by extending current neurophysiological studies on valence processing across sensory domains by incorporating both familiarity and valence into the research design. How core affect is represented in the brain (Barrett & Bliss-Moreau, 2009) and what aspects of valence and arousal processing are domain-specific and which are domain-general, is still a matter of debate (Gao & Shinkareva, 2021; Kiyokawa & Hayashi, 2024; Lin et al., 2020; Marin et al., 2012). The inclusion of the study of dominance, a

third dimension sometimes studied in relation to core affect, may yield further insights into this issue (Jerram et al., 2014; Lang et al., 1997; Sutton et al., 2019).

From a wider perspective, our studies also contribute to the ongoing discussion about whether everyday emotions differ from aesthetic emotions (Fingerhut & Prinz, 2020; Juslin, 2013; Menninghaus et al., 2019; Skov & Nadal, 2020; Skov & Nadal, 2021b; Stamkou et al., 2024), or to put it more simply, whether a special category of emotions, namely aesthetic emotions (Menninghaus et al., 2019; Scherer, 2004; Schindler et al., 2017; Schubert, 2024), are evoked during the perception of aesthetic objects. We focused on core affect, specifically on felt – and not perceived - pleasantness and arousal, and not on discrete emotion categories. Our studies revealed that the emotion spaces of environmental scenes and representational paintings were similar in shape and that for both types of stimuli the correlations between familiarity and pleasantness, arousal and complexity, respectively, were similar in direction. This may suggest that core affect is similarly represented when visual stimuli do not differ regarding the semantic and affective contents but only in their surface structure (Marin & Leder, 2013, 2016). On the other hand, we noted that the pattern of results regarding individual differences (gender and affective personality traits) of complexity and arousal ratings was different for the perception of environmental scenes and representational paintings (Marin & Leder, 2013, 2018), suggesting that person-related factors play a greater role when perceiving visual art than environmental scenes. One can surmise that this difference may have an evolutionary significance enhancing survival.

Menninghaus et al. (2019) propose that discrete aesthetic emotions are associated with felt pleasure and displeasure as well as with arousal, and that pleasure and pleasantness are intrinsic to aesthetic emotions. This is an interesting approach that attempts to merge aspects of two different emotion models into one (discrete emotion models vs. dimensional emotion models) and that also indirectly acknowledges the work by Berlyne and others by integrating concepts such as pleasantness and arousal into their framework of aesthetic emotions. Menninghaus et al. (2019) suggest (see also Schindler et al., 2017), referring to Kant, that the feeling of beauty is a type of emotion that is generated by “superimposing an emotional meaning on prototypical aesthetic virtue terms” (p. 185), and moreover, that aesthetic emotions predict liking and disliking of aesthetic objects. Beauty and liking are frequently regarded as aesthetic judgements rather than aesthetic emotions (Pelowski et al.,

2016). Nevertheless, our behavioural studies demonstrated that concepts such as liking, beauty and pleasantness are indeed distinct (Marin et al., 2016). Menninghaus et al. (2019) also discuss the decisive role of familiarity and novelty in the context of aesthetic emotions, and our findings support the idea that familiarity is another key concept, alongside various types of aesthetic judgements, that should be considered in models of aesthetic experiences. The authors also point out that motivational tendencies of approach and avoidance are characteristic of aesthetic emotions, although these behavioural components of emotion are normally attributed to everyday (utilitarian) emotions. In principle, this suggests that Menninghaus et al. (2019) are not opposed to the idea of considering both motivation and emotion in the study of aesthetic experiences (see Berlyne, 1971; McCrae, 2024).

3.4 Complexity across sensory and artistic domains

Complexity has been of considerable interest to researchers in the field and been frequently studied in the visual domain (for a review see Marin, 2022). For instance, complexity was a variable in early “aesthetic formulae” (Birkhoff, 1933; Eysenck, 1941), and researchers have sought for a long time to predict subjective complexity on the basis of stimulus features (e.g., Attneave, 1957; Berlyne, 1958; Chipman, 1977; Hall, 1969; Ichikawa, 1985; Stenson, 1966). Information-theoretic measures of complexity have become popular mostly since Berlyne’s contribution to the field (for a review see Donderi, 2006; Marin, 2022). Following this trend, this doctoral thesis investigated whether the same set of 13 objective complexity measures, based on edge detection algorithms as well as on information-theoretic measures (for a similar approach see Machado et al., 2015), can predict subjective complexity of environmental scenes and representational paintings (Marin & Leder, 2013). We concluded that the performance of a specific complexity measure depends on the stimulus type, which calls for studying different stimulus categories in greater depth, for example, visual stimuli varying in semantic and affective contents. For instance, compressed file size was a better predictor of subjective complexity in environmental scenes than of paintings, and different types of edge detection algorithms performed well for each of the two stimulus sets. Our objective measures of complexity were unrelated to familiarity (Forsythe et al., 2017) and the relationship between subjective and objective complexity measures did not vary much between presentation durations of 1, 5 and 25 s. Interestingly, we also found that better performances of objec-

tive complexity measures were achieved when subdimension of complexity, and not general complexity ratings, were considered (Marin & Leder, 2016; see also Gartus & Leder, 2017). We also confirmed that the subjective complexity of the representational visual stimuli used in Marin and Leder (2013) was mostly related to the number of elements/objects in a scene and to a lesser degree to the dimension of (dis)organization. Furthermore, our comparative approach revealed that the multidimensionality of visual complexity (Nadal et al., 2010; Van Geert & Wagemans, 2020) in art perception may be less clearly present when paintings are presented for a short duration of 1 s (Marin & Leder, 2016). Studying the effect of presentation duration on the relationship between subjective and objective complexity (in particular short durations under 1 s) as well as acknowledging the multidimensionality of visual complexity may be a fruitful path for the future.

Our computational approach based on the use of a large set of various measures of objective complexity should also be applied in the future because it has been shown that subjective complexity can be better predicted by linear combinations of objective measures than by one single measure (see also Braun et al., 2013; Corchs et al., 2016; Gartus & Leder, 2017; Machado et al., 2015, Redies et al., 2025, Zhou et al., 2023). In most cases, our current knowledge is not advanced enough to justify the choice of one particular measure for one type of stimulus over another, and unfortunately, researchers are often not justifying their particular choice in research reports. However, due to the continuing interest in this topic it should ultimately be possible to obtain a good understanding of how subjective complexity relates to objective stimulus features when viewing aesthetic and non-aesthetic visual stimuli, which in the end will provide a much better basis for predicting evaluative ratings associated with these stimuli as it is currently possible (see, e.g., Hayn-Leichsenring et al., 2017; Iigaya et al., 2021; Nath et al., 2024; Nicki & Moss, 1975; Sidhu et al., 2018). Needless to say, the study of individual differences regarding complexity and evaluative ratings will complement these efforts (Lyssenko et al., 2016; Marin & Leder, 2018; Vissers et al., 2020). Future studies will also have to assess the extent to which our findings are specific to visual stimuli containing semantic and affective contents by employing the same set of objective complexity measures in the analysis of abstract visual stimuli. It has recently been shown that the complexity of a visual scene is determined by both low-level image properties and semantic contents (Kyle-Davidson et al., 2023, Zhou et al., 2023).

In this doctoral thesis, complexity research based on computational methods in the visual domain inspired complexity research in the musical domain (Marin & Leder, 2013), and we have demonstrated that objective measures based on compressed file size and edge detection can be successfully applied to the study of musical complexity. The employed objective measures were relatively easy to compute and performed similarly well as a set of more complex measures based on acoustic sound analyses (Streich, 2007). Therefore, our measures have proven to be useful to obtain a good estimate of stimulus complexity, especially since they were mostly unrelated to (both visual and musical) familiarity (see also Forsythe et al., 2008). The field of music information retrieval has shown large interest in modelling musical complexity, mostly based on acoustic and musical feature analyses (e.g., Berardinis et al., 2022; Giorgi et al., 2017; Weiß & Müller, 2014), but information-theoretic measures have also become popular more recently (Guerrero et al., 2021). For instance, Delplanque et al. (2019) studied the complexity of tone sequences by estimating prediction errors by using entropy. They reported an inverted-U shape relationship between complexity and preference, supporting Berlyne's theory (1971). However, we propose that simple, comprehensive measures, such as compressed file size and event density, are a valid alternative to computationally intensive measures, and that these measures could be useful when working with large stimulus sets and thus have broad applications.

Our experiments also offer insights into the perceptual similarities between musical and visual complexity. In Marin and Leder (2013), we demonstrated by presenting a mixture of two musical genres that excerpts with a higher number of musical instruments (piano trios) were perceived as more complex than excerpts of solo piano music. This auditory manipulation is comparable to simple and complex visual scenes which contain either a low or high number of elements. Future research should study whether the subdimensions of visual complexity, namely the number of elements, (dis)organization and symmetry, can also be perceived when listening to real music, such as Romantic music or other genres. This would be a continuation of the approach followed for visual stimuli in Experiment 3 in Marin and Leder (2016). Our musical stimuli were not selected on the basis of any factors associated with complexity but rather on the affective contents. Therefore, it may be interesting to study whether any "visual" subdimensions of musical complexity can be detected among musical stimuli (analogous to the subdimensions of visual complexity), and

how they may relate to acoustic and musical features. Inspired by the current doctoral thesis proposal, Clemente and Nadal studied liking in response to balance, contour, symmetry and complexity in (artificial) visual designs and melodies and mostly reported domain-specific effects (Clemente et al., 2021). This type of comparative research on musical complexity motivated by findings from the visual domain (see also Clemente et al., 2022; Pease et al., 2018), may be extended to ecological, unaltered stimulus materials in the future. These studies may also further investigate the role of individual differences (Güçlütürk & van Lier, 2019; Marin & Leder, 2018; Van Geert & Wagemans, 2021) as well as the dynamics of complexity perception in greater depth (Marin & Leder, 2016), which is probably only meaningful in the context of a systematic comparison involving various types of visual and musical stimuli.

Our research (Marin et al., 2016) also addressed the question of how subjective complexity across sensory and artistic domains is related to other subjective measures relevant in the study of aesthetic experiences, particularly to familiarity, arousal and different evaluative measures (i.e., liking, beauty and pleasantness) (see also Nath et al., 2024; Van Geert & Wagemans, 2021). We compared three types of visual stimuli varying in their aesthetic quality (environmental scenes, environmental scenes turned into cartoons and representational paintings) and instrumental music. Subjective complexity, mostly determined by the number and type of visual elements in a scene, correlated similarly with our three evaluative measures when perceiving environmental scenes and environmental scenes turned into cartoons, probably due to the fact that liking, beauty and pleasantness correlated very highly with each other. In contrast, the correlations between evaluative measures and complexity considerably varied for representational paintings and music; however, not in the same fashion, which demonstrates domain-specific characteristics that may be due to sensory differences or to the fact that representational paintings contain semantic meaning. At the same time, our data revealed that hedonic tone measures are less correlated with each other when studying artistic stimuli, which may be a potential characteristic for stimuli of aesthetic quality. Alternatively, Che et al. (2024) propose that task instructions may explain why in some studies correlations between evaluative ratings differ and why in others they are very similar to each other (see Table 1 in Che et al., 2024, for a summary of results based on eleven studies mostly using visual stimuli).

Future studies on complexity perception in relation to evaluative judgements may thus assess more than one measure of hedonic tone (see Sidu et al., 2018),

preferably use both within- and between-subjects designs, control for features of complexity (Van Geert & Wagemans, 2021) and employ a wide range of stimuli differing in their aesthetic quality. At the moment, most comparative studies on complexity only focus on one hedonic (evaluative) measure, often without any justifications, for example, there is a current focus on liking and not on beauty (Clemente et al., 2023). Other studies involve both beauty and liking ratings (Che et al., 2024), but do not assess complexity, which makes such studies less informative regarding theory development in the context of aesthetic experiences.

Marin (2015) proposed a comparative approach to neuroaesthetics, and more generally, to empirical aesthetics. Neuroimaging studies have recently focused on the nature of complexity across sensory domains. Güçlütürk et al. (2018) conducted an fMRI-study in which they used stimulus-evoked BOLD responses to musical and visual stimuli to predict the values of the respective objective complexity measures. The complexity measures for music (FLAC, Ogg and event density) were chosen based on the results in Marin and Leder (2013). Several interesting insights emerged, such as the fact that the statistical models for the visual stimuli worked best for brain areas of the visual cortex lower in the processing hierarchy (V1) than for those in the higher visual cortex (V4). In V1, the overlap of the amount of voxels predicted by the three objective complexity measures was considerably high, and this overlap decreased in the higher processing areas. There was a positive correlation between the activation of the voxels representing complexity and neural activity in the lower visual cortex but not in the higher visual cortex. For music, activation of areas sensitive to complexity was stronger in the early auditory cortex and on average lower in the auditory association cortex. Brain activity predicted the compression measures better than event density, which supports the view that compression file measures (i.e., Kolmogorov complexity) are able to capture stimulus complexity effectively across domains (Marin & Leder, 2013). The authors also noted some overlap between areas that have been described in the context of story listening and auditory math tasks. Altogether, the comparative approach followed in our behavioural studies can be taken as a proof of concept that has uncovered several interesting similarities and differences regarding the study of visual and musical complexity, which has clearly inspired other researchers including neuroscientists.

3.5 Nature of aesthetic experiences

For three thousand years, Western philosophers, and later psychologists, have speculated about whether aesthetic experiences are of a special kind (Hayn-Leichsenring & Chatterjee, 2019; Hilgers, 2016; McCrae, 2024; Rind, 2002), differing from everyday perception and in terms of underlying neural responses (Chatterjee & Vartanian, 2014; Karim et al., 2022; Liu et al., 2025; Nadal & Skov, 2025; Skov & Nadal, 2021a; Vessel, 2022; Welter & Lotte, 2024). For instance, is there a qualitative difference between experiencing a real snake on the ground and viewing a painting thereof? Intuitively, one would say yes, arguing that the snake on the ground requires some sort of motivational response, whereas one could enjoy the beauty of the painted snake without any need to act (see Sarasso et al., 2020). However, whether the emotional responses to a real and a painted snake differ (see discussion on aesthetic emotions above and Menninghaus et al., 2019), and whether humans perceive a painted snake in a mental state based on a special cognitive processing mode (i.e., aesthetic attitude, aesthetic stance and disinterested attention, see Basu & Srinivasan, 2024; Bullough, 1912; Palmer et al., 2013) are current questions without any satisfactory answers.

Our comparative experiments were designed to draw some conclusions about the nature of aesthetic experiences (Marin, 2015). Working with the same stimulus material across several studies, while employing different methodologies and presentation durations, we systematically compared subjective and behavioural responses to environmental scenes and representational paintings that were matched for semantic and affective contents as well as for visual complexity (Marin & Leder, 2013, 2016, 2018, 2022; Marin et al., 2016). We were interested in the image-type (bottom-up-factor) that differed between the two categories and not in framing the same image as either art or non-art (top-down factor) (Gerger et al., 2014; Wagner et al., 2014). As discussed in the previous sections, we documented several notable differences between the cognitive and affective processing of environmental scenes and representational paintings. Our results are based on laboratory studies that employed digital images that differ on their surface structure (see differences in objective image features in Marin & Leder, 2013) and thus can easily be classified as art vs. non-art (see also essentialism as discussed in Hayn-Leichsenring, 2017). However, although laboratory studies have great value, aesthetic experiences of real artworks and real environmental scenes in everyday social settings may differ from experiencing digital

images on screens under controlled conditions (see Pelowski, Forster, et al., 2017). In an experiment, in which, for example paintings and environmental scenes were experienced in an ecologically valid situation, one may expect larger absolute values in subjective ratings and behavioural measures such as eye movements than in a laboratory setting, but not necessarily larger relative differences between the picture categories (Brieber et al., 2015; Brieber et al., 2014; Locher et al., 2001). Future studies that combine mobile eye-tracking technology with subjective ratings *in situ* will reveal whether this is indeed the case.

Particularly noteworthy is our eye movement study, in which we employed a free viewing task and two different presentation durations (short vs. long) (Marin & Leder, 2022). For a presentation duration of 25 s, we observed a greater number of fixations for paintings compared to environmental scenes. The main finding, however, was that pleasantness, and not visual complexity, led to opposite effects on eye movement parameters when comparing the two stimulus types: fixation durations were longer for pleasant paintings and shorter for pleasant environmental scenes. Since many researches make the claim that pleasantness and pleasure are characteristic of aesthetic emotions and experiences (Menninghaus, 2019, Skov & Nadal, 2021), the finding that pleasantness led to differential results is particularly meaningful in this context. In this light, our results can also be interpreted as suggestive of a viewing mode (i.e., aesthetic stance) that may be unconsciously adapted when viewing artworks (see also Papenmeier et al., 2024). It is important to note that participants were viewing both types of stimuli in a within-subjects design without any specific viewing or rating task, but of course, after several trials in the experiment participants probably pre-classified all images as either art or non-art. Thus, our findings may hint at a mental state (related to pleasure/pleasantness) that is automatically activated when viewing many artworks in a row (see also Bohrn et al., 2013; Chatterjee et al., 2009). Alternatively, one could interpret our findings as indicative of a state of absorption and flow (Husselman et al., 2024; Kuijpers et al., 2017). One way to clarify this issue in the future is to complement eye movement measurements with electroencephalography (Welke & Vessel, 2022). Current research by Papenmeier et al. (2024) also suggests that comparing different stimulus types is a better way to study this topic than framing the same stimuli either as art or non-art (see also Casteau & Smith, 2024).

Studying the duration and time course of aesthetic experiences is a timely topic (Bachmann & Vipper, 1983; Belfi et al., 2019; Muth et al., 2015; Verhavert et al., 2018; Wagner et al., 2021, see also Leder et al., 2022). We also gained insights into the time course of aesthetic experiences by varying the presentation duration of our stimuli (Marin & Leder, 2013, 2016, 2022; Marin et al., 2016), from relatively short to long (1-25 s). We reported several effects on subjective ratings and eye movement parameters. Future research may extend our approach by employing a larger set of presentation durations, type of rating scales and stimulus types. Based on these findings, the field will also need to reflect on whether, for example, quick beauty judgements of faces and paintings count as aesthetic experiences in a similar way as the beholding of a painting for several minutes. Can the experience be described and modelled as being of the same type? Moreover, when comparing visual with musical aesthetic experiences (or more generally, the experience of static narrativity (visual arts) vs. performing arts (music), Brown, 2022), it may be worthwhile considering differences between sustained aesthetic experiences, such as when listening to a musical piece for several minutes and longer, and a series of aesthetic experiences, such as viewing many different paintings in a row (Khaw & Freedberg, 2018; Rodriguez et al., 2024).

The nature of aesthetic experiences within and across sensory and artistic domains can also be determined by getting a deeper understanding of how subjective experiences of stimulus features are associated with internal state scales related to core affect, discrete aesthetic emotions and aesthetic judgements (see also Russell & George, 1990). By comparing environmental scenes with environmental scenes converted into cartoons and representational paintings, we demonstrated that during the perception of visual art, beauty, liking and pleasantness were differently associated with visual complexity than during the perception of everyday scenes. The specific correlation patterns were different for music. The correlation between stimulus complexity with arousal was the lowest for environmental scenes, followed by the one for representational paintings, and the highest for music, suggesting that during aesthetic experiences complexity and arousal are more strongly correlated. For ratings of environmental scenes and cartoons, beauty, liking and pleasantness were very highly correlated with each other, but for representational paintings these correlations were somewhat lower, which was also the case for music. Beauty is one of the core concepts of aesthetics, and its link to pleasure has been frequently pro-

posed (Brielmann & Pelli, 2019; Skov & Nadal, 2021a). Our data supports this view because we reported strong correlations between beauty and pleasantness for artistic stimuli. In this regard, future research will also have to confirm that terms such as pleasure, pleasantness and pleasingness can be interchangeably used. At the same time, the time course of the relationships between various concepts, such as beauty, pleasantness and liking, should be studied in greater depth.

Marin (2015) advocated a systematic comparative approach to the study of aesthetic experiences (see also Jacobsen & Beudt, 2017). We were able to compare subjective responses to visual art and music, which are among the oldest human art forms (Brown, 2022), and both art forms may be rooted in sexual selection (Marin & Gingras, 2024; Miller, 2001; Varella et al., 2011). In our experiments, working only with affective stimuli, we noted several differences and commonalities between the aesthetic experiences of music and visual art: on the one hand, the arousal-pleasantness space had a different shape, the relationship between complexity and arousal (controlled for familiarity) was much stronger for music than for visual art, and beauty, liking and pleasantness correlated differently with complexity in both artistic domains. Musical expertise predicted musical complexity ratings, but art interest was not predictive of visual complexity ratings. On the other hand, arousal was identified as a mediator in the complexity-hedonic tone relationship in environmental scenes, representational paintings and music. Musical and visual complexity could be predicted by Kolmogorov complexity measures. Affective personality traits predicted both visual and musical arousal. Altogether, these findings suggest that there may be domain-general and domain-specific components that determine aesthetic experiences. To be sure, future research needs to employ different stimulus types, such as other musical genres and abstract art, to replicate and extend our findings before such a claim can be made.

The field of neuraesthetics may also profit from a comparative approach. Luo et al. (2025) argue that models of aesthetic experiences based on the study of visual art put much emphasis on cognitive processing related to artistic styles and concepts (Leder et al., 2004, Pelowski & Akiba, 2011), whereas models based on music-induced aesthetic experiences emphasize the communication of emotions (Brattico, 2019; Brattico & Pearce, 2013). A recent meta-analysis of fMRI-studies examining visual and musical aesthetic experiences supports the view that domain-specific models may be more appropriate than domain-general ones (Luo et al., 2025, but

see Brown et al., 2011). The authors employed several advanced meta-analytical approaches to compare neuronal activity across the two art forms. Interestingly, they could neither identify any overlapping brain regions nor networks. Among the main differences were that motor regions were activated during music perception but not art perception, for which frontal areas were particularly activated. The default-mode network was active during art perception but not during music perception, and the salience network was active during music perception but not during art perception. Luo et al. (2025) argue that visual aesthetic experiences may be characterized by inward processes and introspection (i.e., viewer-driven), whereas music perception engages outward processes (i.e., stimulus-driven). Furthermore, they interpret their results as support for the notion that aesthetic experiences of music are tightly connected to musical emotions, whereas visual aesthetic experiences require higher-order cognitive processes referring to meaning and interpretation. Although the findings and interpretation of this meta-analysis of the neural bases of visual art and music seem convincing, this does not preclude that there is no overlap at the phenomenological level. How the brain generates subjective awareness, (i.e., conscious experiences, *qualia*) is still a matter of debate (Seth & Bayne, 2022). Even if the determinants of beauty in the visual arts (Grinde & Husselman, 2023) differ from those of musical beauty (Fleckenstein et al., 2025), the phenomenological experience of beauty may be similar or the same (Brielmann et al., 2021), regardless of the underlying brain networks.

In light of modern technological advancements, we would encourage researchers to design experiments that acknowledge the social and situational context as well as the participative character of art to a greater extent (see also Brown and Dissanayake, 2009). Highly controlled laboratory work only tells one side of the story about the nature of aesthetic experiences, and theory building would truly profit from a broader methodological approach.

3.6 Conclusions

Humans have created and enjoyed art, including visual art and music, for at least 40 000 years (Aubert et al., 2019; Conard et al., 2009). Since Fechner's work in the second half of the 19th century, psychologists have empirically studied how humans respond to artworks and aesthetic objects. Recent research acknowledges the intricacy of aesthetic experiences and has shown that emotions, especially, are in-

herent to aesthetic responses (Stamkou et al., 2024), which may be due to a social function of emotions in cultural evolution, enabling collective living (Keltner & Oatley, 2022). Current models of aesthetic experiences have mostly focused on one artform (Pelowski et al., 2016), and it is still debated whether domain-specific or domain-general models are more appropriate to model aesthetic experiences. To be sure, almost 150 years after Fechner's *Vorschule der Ästhetik* (1876), the field has developed into an exciting playground that attracts an increasing number of psychologists and neuroscientists alike.

In the context of Berlyne's psychobiological model (1971), a domain-general model grounded in the psychology of motivation, we systematically compared aesthetic responses to visual and musical stimuli of affective contents that also varied in structural complexity and aesthetic (artistic) quality. Our approach comprised the study of subjective responses to various variables, computational analysis of the stimulus features, as well as the study of eye movements. We also considered individual differences and exposure duration in several of our experiments. As such, we have followed a rather comprehensive approach to study aesthetic responses, as suggested by Berlyne (1971) and his *New Experimental Aesthetics*.

In light of our findings, several general conclusions can be drawn. First, we have convincingly demonstrated that it is worthwhile scrutinizing older theories that are still of interest to the field, but at the same time, also face problems regarding replication. Through systematic investigations by considering subjective arousal, a variable that has been frequently neglected by others (Marin, 2022), we were able to offer new insights into the uncertainties surrounding Berlyne's psychobiological model. Some aspects of his motivational theory, notably the complexity-arousal relationship or mediation effects of arousal, were supported and replicated several times in our studies (Marin & Leder, 2013, 2016). Other aspects, such as the famous inverted-U relationship between hedonic tone and a collative variable, were not demonstrated when studying affective stimuli (with one exception for music, see Marin et al., 2016). Therefore, our findings suggest a need for more attention to detail and rigour in experimental research before abandoning older theories completely: they may still contain a kernel of truth, although some aspects may be outdated and need to be revised (Marin, 2022). Eronen and Bringmann (2021), discussing the theory crisis in psychology, argue that there is a lack of robust phenomena. In our experiments, for instance, the positive relationship between structural complexity and subjective

arousal seems to be a robust phenomenon across modalities and stimulus categories, which makes it a good candidate for posing a constraint on a theory.

Second, our series of studies has convincingly demonstrated that a systematic comparative approach to empirical aesthetics (Marin, 2015) involving various types of stimulus categories of two modalities is fruitful and impactful, especially if one is interested in getting a deeper understanding of the nature of aesthetic experiences. To cite just a few examples, we demonstrated a range of similarities and differences between the relationships linking several variables, such as familiarity, complexity, arousal and pleasantness (Marin & Leder, 2013; Marin et al., 2016). Importantly, we showed that musical complexity can be effectively studied by employing objective complexity measures that have been used to study visual complexity. Moreover, we observed differences regarding the effect of pleasantness on eye movements when viewing environmental scenes and representational paintings (Marin & Leder, 2022). Together with insights from the field of neuroaesthetics, a comparative approach will probably lead to faster progress in the field than studies that try to elucidate the nature of aesthetic experiences by focusing on a single artform or type of aesthetic object.

Third, we conclude that clear definitions of psychological concepts are warranted if theoretical advances are to be made. We have demonstrated that umbrella terms, such as hedonic tone (Berlyne, 1971) and complexity, need to be carefully defined in the context of empirical studies (Marin et al., 2016, Marin & Leder, 2016; see also Nadal et al., 2010). Beauty, liking and pleasantness are concepts that were previously found to be highly correlated, but this does not necessarily imply that correlations with another variable (e.g., complexity) are of the same direction and strength. Researchers should also better justify the choice of a specific measure of hedonic tone, or aspect of complexity, instead of another. This information is frequently missing or insufficient, and therefore one cannot easily compare results across studies, as we have seen when comparing studies on the inverted-U relationship between complexity and hedonic tone (Marin, 2022). Conceptual clarification has been named as one challenge in the development of good psychological theories. In the natural sciences, it is common practice to refine concepts and to improve the theoretical framework (Eronen & Bringmann, 2021). The field of empirical aesthetics could benefit from such an approach.

Fourth, our research contributes to the trend of studying and considering individual differences in the research design. Fechner was already aware of individual differences in aesthetic responses, and the field has witnessed a recent interest in this topic (e.g., Chen et al., 2022; Gingras et al., 2015; Hogan, 2015; Pombo & Pelli, 2022). We considered individual differences in our statistical models and reported either effects of familiarity, personality, gender or art interest/expertise. The general picture emerging from our results is that individual differences should be studied in relation to one variable of interest as well as regarding the relationships between various variables in the future. When comparing smaller groups of participants it may be worth controlling for certain background variables. In light of these findings, one can argue that the role of individual differences when developing theories of aesthetic experiences should receive more attention, especially if reported effects are replicable. More generally, individual differences may have moderating effects regarding cognitive and affective processing and its interplay, which creates aesthetic experiences. This makes it a demanding and complex topic for future research.

Fifth, several of our studies address the dynamic nature of aesthetic experiences in the visual domain. An aesthetic experience is commonly considered as evolving over time, especially in the context of theories that are based on various stages of information processing (Pelowski, Markey, et al., 2017). For example, we have demonstrated that the multidimensionality of complexity may not be clearly present when representational paintings are shown for 1 s in comparison to 5 s, and that the relationship between subjective and objective measures of complexity may partly depend on the presentation duration (Marin & Leder, 2016). Furthermore, we reported that interrelationships between variables may differ across short vs. long presentation durations, as may be the case for factors affecting eye movement parameters (Marin & Leder, 2022). Thus, we would encourage researchers to examine dynamic aspects of aesthetic experiences to a greater extent as has previously been the case. Time-sensitive methods such as electroencephalography are ideal to complement eye movement measurements and subjective responses. This will ultimately answer how aesthetic experiences emerge and decay in a stream of ordinary, baseline subjective experiences.

Last, we conclude form our studies that there is great value in employing the same stimulus material across several studies, which focus on different research questions related to aesthetic experiences. We were able to replicate and extend

several of our findings (see Shrout & Rodgers, 2018), which suggests that these findings are robust for these stimulus types. Although a repetitive use of stimuli precludes generalization of our results (also because we mostly tested female participants), we can at least make the claim that our findings do not seem to be the result of chance. Future research could extend our comparative approach by employing different visual and auditory stimulus types (e.g., abstract visual stimuli, film, other types of musical genres, environmental sounds), or even expand to other sensory modalities, such as olfactory (e.g., Diessner et al., 2021) and tactile modalities (e.g., Carbon & Jakesch, 2012). There are numerous meaningful ways of systematically comparing responses to different stimulus categories, from which theory building may profit in the long run (Marin, 2015).

4. References

Actis-Grosso, R., Lega, C., Zani, A., Daneyko, O., Cattaneo, Z., & Zavagno, D. (2017). Can music be figurative? Exploring the possibility of crossmodal similarities between music and visual arts. *Psihologija*, *50*(3), 285-306. https://doi.org/10.2298/PSI1703285A

Althuizen, N. (2021). Revisiting Berlyne's inverted U-shape relationship between complexity and liking: The role of effort, arousal, and status in the appreciation of product design aesthetics. *Psychology and Marketing*, *38*(3), 481-503. https://doi.org/10.1002/mar.21449

Anderson, M. D., Elder, J. H., Graf, E. W., & Adams, W. J. (2022). The time-course of real-world scene perception: Spatial and semantic processing. *iScience*, *25*(12), Article 105633. https://doi.org/10.1016/j.isci.2022.105633

Arnheim, R. (1988). *The power of the center: A study of composition in the visual arts: The new version*. University of California Press.

Aslanoglu, R. K., & Olguntürk, N. (2019). Color and visual complexity in abstract images: Part II. *Color Research and Application*, *44*(6), 941-947. https://doi.org/10.1002/col.22408

Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, *53*(4), 221-227. https://doi.org/10.1037/h0043921

Aubert, M., Lebe, R., Oktaviana, A. A., Tang, M., Burhan, B., Hamrullah, Jusdi, A., Abdullah, Hakim, B., Zhao, J. X., Geria, I. M., Sulistyarto, P. H., Sardi, R., & Brumm, A. (2019). Earliest hunting scene in prehistoric art. *Nature*, *576*(7787), 442-445. https://doi.org/10.1038/s41586-019-1806-y

Augustin, M. D., & Leder, H. (2006). Art expertise: A study of concepts and conceptual spaces. *Psychology Science*, *48*(2), 135-156.

Augustin, M. D., Carbon, C.-C., & Wagemans, J. (2012). Artful terms: A study on aesthetic word usage for visual art versus film and music. *i-Perception*, *3*(5), 319-337. https://doi.org/10.1068/i0511aap

Augustin, M. D., Defranceschi, B., Fuchs, H. K., Carbon, C.-C., & Hutzler, F. (2011). The neural time course of art perception: An ERP study on the processing of style versus content in art. *Neuropsychologia*, *49*(7), 2071-2081. https://doi.org/10.1016/j.neuropsychologia.2011.03.038

Bachmann, T., & Vipper, K. (1983). Perceptual rating of paintings from different artistic styles as a function of semantic differential scales and exposure time. *Archiv für Psychologie*, *135*(2), 149-161.

Baltissen, R., & Ostermann, B.-M. (1998). Are the dimensions underlying aesthetic and affective judgment the same? *Empirical Studies of the Arts*, *16*(2), 97-113. https://doi.org/10.2190/KK5M-GAM9-FFFG-0BCC

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-evoked nostalgia: Affect, memory, and personality. *Emotion*, *10*(3), 390-403. https://doi.org/10.1037/a0019006

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*(3), 628-647. https://doi.org/10.1037/0033-295X.113.3.628

Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B*, *364*(1521), 1325-1334. https://doi.org/10.1098/rstb.2008.0312

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167-218. https://doi.org/10.1016/S0065-2601(08)00404-8

Basu, S., & Srinivasan, N. (2024). Disinterested attention and aesthetic experience. *Progress in Brain Research*, *287*, 25-44. https://doi.org/10.1016/bs.pbr.2024.05.012

Beermann, U., Hosoya, G., Schindler, I., Scherer, K. R., Eid, M., Wagner, V., & Menninghaus, W. (2021). Dimensions and clusters of aesthetic emotions: A semantic profile analysis. *Frontiers in Psychology*, *12*, Article 667173. https://doi.org/10.3389/fpsyg.2021.667173

Belfi, A. M., Vessel, E. A., Brielmann, A., Isik, A. I., Chatterjee, A., Leder, H., Pelli, D. G., & Starr, G. G. (2019). Dynamics of aesthetic experience are reflected in the default-mode network. *NeuroImage*, *188*, 584-597. https://doi.org/10.1016/j.neuroimage.2018.12.017

Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A., & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications*, *8*(1), Article 42. https://doi.org/10.1186/s41235-023-00499-6

Berardinis, J. d., Cangelosi, A., & Coutinho, E. (2022). Measuring the structural complexity of music: From structural segmentations to the automatic evaluation of models for music generation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *30*, 1963-1976. https://doi.org/10.1109/TASLP.2022.3178203

Berlyne, D. E. (1958). The influence of complexity and novelty in visual figures on orienting responses. *Journal of Experimental Psychology*, *55*(3), 289-296. https://doi.org/10.1037/h0043555

Berlyne, D. E., Ogilvie, J., & Parham, L. (1968). The dimensionality of visual complexity, interestingness, and pleasingness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *22*(5), 376-387. https://doi.org/10.1037/h0082777

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill Book Company. https://doi.org/10.1037/11164-000

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*(5), 279-286. https://doi.org/10.3758/BF03212593

Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. Appleton-Century-Crofts.

Berlyne, D. E. (Ed.) (1974). *Studies in the New Experimental Aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Hemisphere.

Berridge C. W. (2008). Noradrenergic modulation of arousal. *Brain Research Reviews*, *58*(1), 1–17. https://doi.org/10.1016/j.brainresrev.2007.10.013

Bhandari, U., Chang, K., & Neben, T. (2019). Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. *Information & Management*, *56*(1), 85-93. https://doi.org/10.1016/j.im.2018.07.003

Bigand, E., Filipic, S., & Lalitte, P. (2005). The time course of emotional responses to music. *Annals of the New York Academy of Sciences*, *1060*(1), 429-437. https://doi.org/10.1196/annals.1360.036

Birkhoff, G. D. (1933). *Aesthetic measure*. Harvard University Press. https://doi.org/10.4159/harvard.9780674734470

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research*, *49*(24), 2992-3000. https://doi.org/10.1016/j.visres.2009.09.014

Blacking, J. (1973). *How musical is man?* University of Washington Press.
Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences USA*, *98*(20), 11818-11823. https://doi.org/10.1073/pnas.191355898
Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2013). When we like what we know – A parametric fMRI analysis of beauty and familiarity. *Brain and Language*, *124*(1), 1-8. https://doi.org/10.1016/j.bandl.2012.10.003
Boon, J. P., Casti, J., & Taylor, R. P. (2011). Artistic forms and complexity. *Nonlinear Dynamics Psychology and Life Sciences*, *15*(2), 265-283.
Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, *16*(4), 756-766. https://doi.org/10.1177/1745691620969647
Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, *44*(3), 364-373. https://doi.org/10.1111/j.1469-8986.2007.00520.x
Bradley, M. M., Houbova, P., Miccoli, L., Costa, V. D., & Lang, P. J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, *48*(11), 1544-1553. https://doi.org/10.1111/j.1469-8986.2011.01223.x
Brattico, E. (2019). The neuroaesthetics of music: A research agenda coming of age. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford Handbook of Music and the Brain* (pp. 364-390). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198804123.013.15
Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology*, *4*, Article 206. https://doi.org/10.3389/fpsyg.2013.00206
Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 48-61. https://doi.org/10.1037/a0031624
Bravo, F., Glogowski, J., Stamatakis, E. A., & Herfert, K. (2024). Dissonant music engages early visual processing. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(30), Article e2320378121. https://doi.org/10.1073/pnas.2320378121
Braun, J., Amirshahi, S. A., Denzler, J., & Redies, C. (2013). Statistical image properties of print advertisements, visual artworks and images of architecture. *Frontiers in Psychology*, *4,* Article 808. https://doi.org/10.3389/fpsyg.2013.00808
Brieber, D., Forster, M., & Leder, H. (2020). On the mutual relation between art experience and viewing time. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(2), 197-208. https://doi.org/10.1037/aca0000210
Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica*, *154*, 36-42. https://doi.org/10.1016/j.actpsy.2014.11.004
Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in time and space: Context modulates the relation between art experience and viewing time. *PLOS One*, *9*(6), Article e99019. https://doi.org/10.1371/journal.pone.0099019
Brielmann, A. A., & Dayan, P. (2022). A computational model of aesthetic value. *Psychological Review*, *129*(6), 1319-1337. https://doi.org/10.1037/rev0000337

Brielmann, A. A., Nuzzo, A., & Pelli, D. G. (2021). Beauty, the feeling. *Acta Psychologica*, *219*, Article 103365. https://doi.org/10.1016/j.actpsy.2021.103365

Brielmann, A. A., & Pelli, D. G. (2019). Intense beauty requires intense pleasure. *Frontiers in Psychology*, *10*, Article 2420. https://doi.org/10.3389/fpsyg.2019.02420

Brielmann, A. A., Vale, L., & Pelli, D. G. (2017). Beauty at a glance: The feeling of beauty and the amplitude of pleasure are independent of stimulus duration. *Journal of Vision*, *17*(14), 9-9. https://doi.org/10.1167/17.14.9

Brown, S. (2022). *The unification of the arts: A framework for understanding what the arts share and why*. Oxford University Press.

Brown, S., & Dissanayake, E. (2009). The arts are more than aesthetics: Neuroaesthetics as narrow aesthetics. In M. Skov & O. Vartanian (Eds.), *Neuroaesthetics* (pp. 43-57). Baywood.

Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, *58*(1), 250-258. https://doi.org/10.1016/j.neuroimage.2011.06.012

Bruns, A., Pombo, M., Ripollés, P., & Pelli, D. G. (2023). Emotions of subject and object affect beauty differently for images and music. *Journal of Vision*, *23*(13), Article 6. https://doi.org/10.1167/jov.23.13.6

Bullack, A., Budenbender, N., Roden, I., & Kreutz, G. (2018). Psychophysiological responses to "happy"' and "sad" music: A replication study. *Music Perception*, *35*(4), 502-517. https://doi.org/10.1525/mp.2018.35.4.502

Bullot, N. J., & Reber, R. (2013). The artful mind meets art history: Toward a psycho-historical framework for the science of art appreciation. *Behavioral and Brain Sciences*, *36*(2), 123-137. https://doi.org/10.1017/S0140525X12000489

Bullough, E. (1912). "Psychical Distance' as a factor in art and an aesthetic principle. *British Journal of Psychology*, *5*, 87-118. https://doi.org/10.1111/j.2044-8295.1912.tb00057.x

Campos-Bueno, J. J., DeJuan-Ayala, O., Montoya, P., & Birbaumer, N. (2015). Emotional dimensions of music and painting and their interaction. *Spanish Journal of Psychology*, *18*, Article E54. https://doi.org/10.1017/sjp.2015.53

Carballal, A., Fernandez-Lozano, C., Rodriguez-Fernandez, N., Santos, I., & Romero, J. (2020). Comparison of outlier-tolerant models for measuring visual complexity. *Entropy*, *22*(4), Article 488. https://www.mdpi.com/1099-4300/22/4/488

Carballal, A., Santos, A., Romero, J., Machado, P., Correia, J., & Castro, L. (2018). Distinguishing paintings from photographs by complexity estimates. *Neural Computing and Applications*, *30*(6), 1957-1969. https://doi.org/10.1007/s00521-016-2787-5

Carbon, C. C., & Jakesch, M. (2012). A model for haptic aesthetic processing and its implications for design. *Proceedings of the IEEE*, *101*(9), 2123-2133. https://doi.org/10.1109/JPROC.2012.2219831

Casteau, S., & Smith, D. T. (2024). How does contextual information affect aesthetic appreciation and gaze behavior in figurative and abstract artwork? *Journal of Vision*, *24*(12), Article 8. https://doi.org/10.1167/jov.24.12.8

Celikors, E., & Field, D. J. (2025). Beauty is in the eye of your cohort: Structured individual differences allow predictions of individualized aesthetic ratings of images. *Cognition*, *256*, Article 106036. https://doi.org/10.1016/j.cognition.2024.106036

Cespedes-Guevara, J., & Eerola, T. (2018). Music communicates affects, not basic emotions – A constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology*, *9,* Article 215. https://doi.org/10.3389/fpsyg.2018.00215

Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response to facial attractiveness. *Neuropsychology*, *23*(2), 135-143. https://doi.org/10.1037/a0014430

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences*, *18*(7), 370-375. https://doi.org/10.1016/j.tics.2014.03.003

Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, *1369*(1), 172-194. https://doi.org/10.1111/nyas.13035

Che, J., Sun, X., Skov, M., & Nadal, M. (2024). Judging beauty and liking: The effects of personal intuitions and task instructions. *Psychology of Aesthetics, Creativity, and the Arts*, Advance online publication. https://doi.org/10.1037/aca0000702

Chen, Y.-C., Chang, A., Rosenberg, M. D., Feng, D., Scholl, B. J., & Trainor, L. J. (2022). “Taste typicality” is a foundational and multi-modal dimension of ordinary aesthetic experience. *Current Biology*, *32*(8), 1837-1842. https://doi.org/10.1016/j.cub.2022.02.039

Chipman, S. F. (1977). Complexity and structure in visual patterns. *Journal of Experimental Psychology: General*, *106*(3), 269-301. https://doi.org/10.1037/0096-3445.106.3.269

Chmiel, A., & Schubert, E. (2018). Emptying rooms: When the inverted-U model of preference fails - An investigation using music with collative extremes. *Empirical Studies of the Arts*, *36*(2), 199-221. https://doi.org/10.1177/0276237417732683

Chmiel, A., & Schubert, E. (2019). Unusualness as a predictor of music preference. *Musicae Scientiae*, *23*(4), 426-441. https://doi.org/10.1177/1029864917752545

Christensen, J. F., & Calvo-Merino, B. (2013). Dance as a subject for empirical aesthetics. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 76-88. https://doi.org/10.1037/a0031827

Christensen, J. F., Gaigg, S. B., Gomila, A., Oke, P., & Calvo-Merino, B. (2014). Enhancing emotional experiences to dance through music: The role of valence and arousal in the cross-modal bias. *Frontiers in Human Neuroscience*, *8*, Article 757. https://doi.org/10.3389/fnhum.2014.00757

Clemente, A., Friberg, A., & Holzapfel, A. (2023). Relations between perceived affect and liking for melodies and visual designs. *Emotion*, *23*(6), 1584-1605. https://doi.org/10.1037/emo0001141

Clemente, A., Pearce, M. T., & Nadal, M. (2022). Musical aesthetic sensitivity. *Psychology of Aesthetics, Creativity, and the Arts*, *16*(1), 58-73. https://doi.org/10.1037/aca0000381

Clemente, A., Pearce, M. T., Skov, M., & Nadal, M. (2021). Evaluative judgment across domains: Liking balance, contour, symmetry and complexity in melodies and visual designs. *Brain and Cognition*, *151*, Article 105729. https://doi.org/10.1016/j.bandc.2021.105729

Cleridou, K., & Furnham, A. (2014). Personality correlates of aesthetic preferences for art, architecture, and music. *Empirical Studies of the Arts*, *32*(2), 231-255. https://doi.org/10.2190/EM.32.2.f

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*(4), 731-738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Colombetti, G., & Kuppens, P. (2024). How should we understand valence, arousal, and their relation? In A. Scarantino (Ed.), *Emotion theory: The Routledge comprehensive guide* (pp. 599-620). Routledge.

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, *460*(7256), 737-740. https://doi.org/10.1038/nature08169

Corchs, S. E., Ciocca, G., Bricolo, E., & Gasparini, F. (2016). Predicting complexity perception of real world images. *PLOS One*, *11*(6), Article e0157986. https://doi.org/10.1371/journal.pone.0157986

Cotter, K. N., Fekete, A., & Silvia, P. J. (2022). Why do people visit art museums? Examining visitor motivations and visit outcomes. *Empirical Studies of the Arts*, *40*(2), 275-295. https://doi.org/10.1177/02762374211011740

Coursaris, C. K., Swierenga, S. J., & Watrall, E. (2008). An empirical investigation of color temperature and gender effects on web aesthetics. *Journal of Usability Studies*, *3*(3), 103-117.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Crozier, J. B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the New Experimental Aesthetics: Steps toward an objective psychology of aesthetic appreciation* (pp. 27-90). Hemisphere.

Cupchik, G. C. (1988). The legacy of Daniel E. Berlyne. *Empirical Studies of the Arts*, *6*(2), 171-186. https://doi.org/10.2190/flm8-6nq7-n5wm-wllt

Cupchik, G. C. (2002). The evolution of psychical distance as an aesthetic concept. *Culture & Psychology*, *8*(2), 155-187. https://doi.org/10.1177/1354067x02008002437

Cupchik, G. C. (2016). *The aesthetics of emotion: Up the down staircase of the mind-body* (Studies in Emotion and Social Interaction). Cambridge University Press. https://doi.org/10.1017/CBO9781139169301

Cupchik, G. C. (2022). One hundred and fifty years after Fechner: A view from the "Middle of the Storm". *Psychology of Aesthetics Creativity and the Arts*, *16*(3), 544-552. https://doi.org/10.1037/aca0000352

Cupchik, G. C., & Berlyne, D. E. (1979). The perception of collative properties in visual stimuli. *Scandinavian Journal of Psychology*, *20*(1), 93-104. https://doi.org/10.1111/j.1467-9450.1979.tb00688.x

Cupchik, G. C., & Gebotys, R. J. (1988). The experience of time, pleasure, and interest during aesthetic episodes. *Empirical Studies of the Arts*, *6*(1), 1-12. https://doi.org/10.2190/5YN3-J3P8-FWHY-UDB3

Curwen, C. (2018). Music-colour synaesthesia: Concept, context and qualia. *Consciousness and Cognition*, *61*, 94-106. https://doi.org/10.1016/j.concog.2018.04.005

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95-111. https://doi.org/10.1016/S0301-0511(99)00044-7

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2023). Aesthetic and physiological effects of naturalistic multimodal music listening. *Cognition*, *239*, Article 105537. https://doi.org/10.1016/j.cognition.2023.105537

Damiano, C., Wilder, J., Zhou, E. Y., Walther, D. B., & Wagemans, J. (2023). The role of local and global symmetry in pleasure, interest, and complexity judgments of natural scenes. *Psychology of Aesthetics, Creativity, and the Arts, 17*(3), 322-337. https://doi.org/10.1037/aca0000398

Damm, L., Varoqui, D., De Cock, V. C., Dalla Bella, S., & Bardy, B. (2020). Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. *Neuroscience & Biobehavioral Reviews*, *112*, 553-584. https://doi.org/10.1016/j.neubiorev.2019.12.024

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*(2), 468-477. https://doi.org/10.3758/s13428-011-0064-1

Davies, S. (2012). *The artful species: Aesthetics, art, and evolution*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199658541.001.0001

Day, H. (1967). Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. *Perception & Psychophysics*, *2*(7), 281-286. https://doi.org/10.3758/BF03211042

de Cruys, S. V., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *i-Perception*, *2*(9), 1035-1062. https://doi.org/10.1068/i0466aap

de Fleurian, R., Blackwell, T., Ben-Tal, O., & Müllensiefen, D. (2017). Information-theoretic measures predict the human judgment of rhythm complexity. *Cognitive Science*, *41*(3), 800-813. https://doi.org/10.1111/cogs.12347

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences, 7*(10), 460-467. https://doi.org/10.1016/j.tics.2003.08.014

Delplanque, J., De Loof, E., Janssens, C., & Verguts, T. (2019). The sound of beauty: How complexity determines aesthetic preference. *Acta Psychologica*, *192*, 146-152. https://doi.org/10.1016/j.actpsy.2018.11.011

Demmer, T. R., Kühnapfel, C., Fingerhut, J., & Pelowski, M. (2023). Does an emotional connection to art really require a human artist? Emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. *Computers in Human Behavior*, *148*, Article 107875. https://doi.org/10.1016/j.chb.2023.107875

Diessner, R., Genthôs, R., Arthur, K., Adkins, B., & Pohling, R. (2021). Olfactory and gustatory beauty: Aesthetic emotions and trait appreciation of beauty. *Psychology of Aesthetics, Creativity, and the Arts, 15*(1), 38–50. https://doi.org/10.1037/aca0000262

Dissanayake, E. (2019). What art is and what art does: An overview of contemporary evolutionary hypotheses. In C. Martindale, P. Locher, V. Petrov, & A. Berleant (Eds.), *Evolutionary and neurocognitive approaches to aesthetics, creativity and the arts* (pp. 1-14). Routledge.

Dolcos, F., Katsumi, Y., Moore, M., Berggren, N., de Gelder, B., Derakshan, N., Hamm, A. O., Koster, E. H. W., Ladouceur, C. D., Okon-Singer, H., Pegna, A. J., Richter, T., Schweizer, S., Van den Stock, J., Ventura-Bort, C., Weymar, M., & Dolcos, S. (2020). Neural correlates of emotion-attention interactions: From perception, learning, and memory to social cognition, individual differ-

ences, and training interventions. *Neuroscience & Biobehavioral Reviews*, *108*, 559-601. https://doi.org/10.1016/j.neubiorev.2019.08.017

Donderi, D. C. (2006). Visual complexity: A review. *Psychological Bulletin*, *132*(1), 73-97. https://doi.org/10.1037/0033-2909.132.1.73

Durmus, D. (2020). Spatial frequency and the performance of image-based visual complexity metrics. *IEEE Access*, *8*, 100111-100119. https://doi.org/10.1109/ACCESS.2020.2998292

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18-49. https://doi.org/10.1177/0305735610362821

Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, *30*(3), 307-340. https://doi.org/10.1525/mp.2012.30.3.307

Eisenman, R. (1966). Pleasing and interesting visual complexity: Support for Berlyne. *Perceptual and Motor Skills*, *23*(3_suppl), 1167-1170. https://doi.org/10.2466/pms.1966.23.3f.1167

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572–595). Oxford University Press.

Ercegovac, I. R., Dobrota, S., & Kuščević, D. (2015). Relationship between music and visual art preferences and some personality traits. *Empirical Studies of the Arts*, *33*(2), 207-227. https://doi.org/10.1177/0276237415597390

Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*, *16*(4), 779-788. https://doi.org/10.1177/1745691620970586

Etani, T., Miura, A., Kawase, S., Fujii, S., Keller, P. E., Vuust, P., & Kudo, K. (2024). A review of psychological and neuroscientific research on musical groove. *Neuroscience & Biobehavioral Reviews*, *158*, Article 105522. https://doi.org/10.1016/j.neubiorev.2023.105522

Eysenck, H. J. (1941). The empirical determination of an aesthetic formula. *Psychological Review*, *48*(1), 83-92. https://doi.org/10.1037/h0062483

Eysenck, H. J. (1942). The experimental study of the 'good Gestalt' - A new approach. *Psychological Review*, *49*(4), 344-364. https://doi.org/10.1037/h0057013

Fan, S., Koenig, B. L., Zhao, Q., & Kankanhalli, M. S. (2020). A deeper look at human visual perception of images. *SN Computer Science*, *1*(1), Article 58. https://doi.org/10.1007/s42979-019-0061-5

Fekete, A., Pelowski, M., Specker, E., Brieber, D., Rosenberg, R., & Leder, H. (2023). The Vienna Art Picture System (VAPS): A data set of 999 paintings and subjective ratings for art and aesthetics research. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(5), 660-671. https://doi.org/10.1037/aca0000460

Fernandez-Lozano, C., Carballal, A., Machado, P., Santos, A., & Romero, J. (2019). Visual complexity modelling based on image features fusion of multiple kernels. *PeerJ*, *7*, Article e7075. https://doi.org/10.7717/peerj.7075

Ferreira-Santos, F. (2016). The role of arousal in predictive coding. *Behavioral and Brain Sciences*, *39*, Article e207. https://doi.org/10.1017/s0140525x15001788

Fingerhut, J., & Prinz, J. J. (2020). Aesthetic emotions reconsidered. *The Monist*, *103*(2), 223-239. https://doi.org/10.1093/monist/onz037

Fleckenstein, A. M., Vuoskoski, J. K., & Dibben, N. (2025). Understanding musical beauty. *Empirical Studies of the Arts*, *43*(1), 505-523. https://doi.org/10.1177/02762374241253771

Forster, M., Leder, H., & Ansorge, U. (2016). Exploring the subjective feeling of fluency. *Experimental Psychology*, *63*(1), 45-58. https://doi.org/10.1027/1618-3169/a000311

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, *40*(1), 116-129. https://doi.org/10.3758/BRM.40.1.116

Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, *102*, 49-70. https://doi.org/10.1348/000712610x498958

Forsythe, A., Street, N., & Helmy, M. (2017). Revisiting Rossion and Pourtois with new ratings for automated complexity, familiarity, beauty, and encounter. *Behavior Research Methods*, *49*(4), 1484-1493. https://doi.org/10.3758/s13428-016-0808-z

Frank, H. (1959). *Grundlagenprobleme der Informationsästhtetik und erste Anwendung auf die mime pure*. Schnelle.

Frascaroli, J., Leder, H., Brattico, E., & Van de Cruys, S. (2024). Aesthetics and predictive processing: Grounds and prospects of a fruitful encounter. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1895), Article 20220410. https://doi.org/doi:10.1098/rstb.2022.0410

Freitas, C., Manzato, E., Burini, A., Taylor, M. J., Lerch, J. P., & Anagnostou, E. (2018). Neural correlates of familiarity in music listening: A systematic review and a neuroimaging meta-analysis. *Frontiers in Neuroscience*, *12,* Article 686. https://doi.org/10.3389/fnins.2018.00686

Friedenberg, J., Martin, P., Uy, N., & Kvapil, M. (2022). Judged beauty of fractal symmetries. *Empirical Studies of the Arts*, *40*(1), 100-120. https://doi.org/10.1177/0276237421994699

Ganczarek, J., Pietras, K., & Rosiek, R. (2020). Perceived cognitive challenge predicts eye movements while viewing contemporary paintings. *PsyCh Journal*, *9*(4), 490-506. https://doi.org/10.1002/pchj.365

Gao, C., & Shinkareva, S. V. (2021). Modality-general and modality-specific audiovisual valence processing. *Cortex*, *138*, 127-137. https://doi.org/10.1016/j.cortex.2021.01.022

Gartus, A., & Leder, H. (2017). Predicting perceived visual complexity of abstract patterns using computational measures: The influence of mirror symmetry on complexity perception. *PLOS One*, *12*(11), Article e0185276. https://doi.org/10.1371/journal.pone.0185276

Gerdes, A. B., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, *5*, Article 1351. https://doi.org/10.3389/fpsyg.2014.01351

Gerger, G., Leder, H., & Kremer, A. (2014). Context effects on emotional and aesthetic evaluations of artworks and IAPS pictures. *Acta Psychologica*, *151*, 174-183. https://doi.org/10.1016/j.actpsy.2014.06.008

Gerger, G., Pelowski, M., & Leder, H. (2018). Empathy, Einfühlung, and aesthetic experience: The effect of emotion contagion on appreciation of representational and abstract art using fEMG and SCR. *Cognitive Processing*, *19*(2), 147-165. https://doi.org/10.1007/s10339-017-0800-2

Gingras, B., Marin, M. M., & Fitch, W. T. (2014). Beyond intensity: Spectral features effectively predict music-induced subjective arousal. *Quarterly Journal of Experimental Psychology*, *67*(7), 1428-1446. https://doi.org/10.1080/17470218.2013.863954

Gingras, B., Marin, M. M., Puig-Waldmüller, E. & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in Human Neuroscience: Cognitive Neuroscience*, *9*, Article 916. https://doi.org/10.3389/fnhum.2015.00619

Giorgi, B. D., Dixon, S., Zanoni, M., & Sarti, A. (2017). A data-driven model of tonal chord sequence complexity. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *25*(11), 2237-2250. https://doi.org/10.1109/TASLP.2017.2756443

Godfrey-Smith, P. (1996). *Complexity and the function of mind in nature*. Cambridge University Press. https://doi.org/10.1017/CBO9781139172714

Goller, J., Mitrovic, A., & Leder, H. (2019). Effects of liking on visual attention in faces and paintings. *Acta Psychologica*, *197*, 115-123. https://doi.org/10.1016/j.actpsy.2019.05.008

Graf, L. K., & Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking. *Personality and Social Psychology Review*, *19*(4), 395-410. https://doi.org/10.1177/1088868315574978

Gregorians, L., Velasco, P. F., Zisch, F., & Spiers, H. J. (2022). Architectural experience: Clarifying its central components and their relation to core affect with a set of first-person-view videos. *Journal of Environmental Psychology*, *82*, Article 101841. https://doi.org/10.1016/j.jenvp.2022.101841

Grekow, J. (2016). Music emotion maps in arousal-valence space. In K. Saeed & W. Homenda (Eds.), *Proceedings of the 15th International Conference on Computer Information Systems and Industrial Management. Lecture Notes in Computer Science* (Vol. 9842, pp. 697-706). Springer. https://doi.org/10.1007/978-3-319-45378-1_60

Grinde, B., & Husselman, T.-A. (2023). An attempt to explain visual aesthetic appreciation. *Integrative Psychological and Behavioral Science*, *57*(3), 840-855. https://doi.org/10.1007/s12124-022-09701-8

Gu, J., Cao, L., & Liu, B. (2019). Modality-general representations of valences perceived from visual and auditory modalities. *NeuroImage*, *203*, Article 116199. https://doi.org/10.1016/j.neuroimage.2019.116199

Güçlütürk, Y., Güçlü, U., van Gerven, M., & van Lier, R. (2018). Representations of naturalistic stimulus complexity in early and associative visual and auditory cortices. *Scientific Reports*, *8*(1), Article 3439. https://doi.org/10.1038/s41598-018-21636-y

Güçlütürk, Y., Jacobs, R. H. A. H., & Lier, R. v. (2016). Liking versus complexity: Decomposing the inverted U-curve. *Frontiers in Human Neuroscience*, *10*, Article 112. https://doi.org/10.3389/fnhum.2016.00112

Güçlütürk, Y., & van Lier, R. (2019). Decomposing complexity preferences for music. *Frontiers in Psychology*, *10,* Article 674. https://doi.org/10.3389/fpsyg.2019.00674

Guerrero, D., Rivera, P., Febres, G., & Gershenson, C. (2021). Towards a measure for characterizing the informational content of audio signals and the relation between complexity and auditory encoding. *Entropy*, *23*(12), Article 1613. https://www.mdpi.com/1099-4300/23/12/1613

Hagerhall, C. M., Laike, T., Taylor, R. P., Küller, M., Küller, R., & Martin, T. P. (2008). Investigations of human EEG response to viewing fractal patterns. *Perception*, *37*(10), 1488-1494. https://doi.org/10.1068/p5918

Haj-Ali, H., Anderson, A. K., & Kron, A. (2020). Comparing three models of arousal in the human brain. *Social Cognitive and Affective Neuroscience*, *15*(1), 1-11. https://doi.org/10.1093/scan/nsaa012

Hall, A. C. (1969). Measures of the complexity of random black and white and coloured stimuli. *Perception and Motor Skills*, *29*(3), 773-774. https://doi.org/10.2466/pms.1969.29.3.773

Hare, F. G. (1974). Artistic training and responses to visual and auditory patterns varying in uncertainty. In Berlyne, E. D. (Ed.), *Studies in the New Experimental Aesthetics: Steps towards an objective psychology of aesthetic appreciation* (pp. 159-168). Hemisphere.

Hashim, S., Stewart, L., Küssner, M. B., & Omigie, D. (2023). Music listening evokes story-like visual imagery with both idiosyncratic and shared content. *PLOS One*, *18*(10), Article e0293412. https://doi.org/10.1371/journal.pone.0293412

Hayn-Leichsenring, G. U. (2017). The ambiguity of artworks – A guideline for empirical aesthetics research with artworks as stimuli. *Frontiers in Psychology*, *8*, Article 1857. https://doi.org/10.3389/fpsyg.2017.01857

Hayn-Leichsenring, G. U., & Chatterjee, A. (2019). Colliding terminological systems - Immanuel Kant and contemporary empirical aesthetics. *Empirical Studies of the Arts*, *37*(2), 197-219. https://doi.org/10.1177/0276237418818635

Hayn-Leichsenring, G. U., Lehmann, T., & Redies, C. (2017). Subjective ratings of beauty and aesthetics: Correlations with statistical image properties in Western oil paintings. *i-Perception*, *8*(3), Article 2041669517715474. https://doi.org/10.1177/2041669517715474

Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405-439. https://doi.org/10.1016/j.pneurobio.2007.09.001

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498-504. https://doi.org/10.1016/j.tics.2003.09.006

Henshilwood, C. S., d'Errico, F., van Niekerk, K. L., Coquinot, Y., Jacobs, Z., Lauritzen, S.-E., Menu, M., & García-Moreno, R. (2011). A 100,000-year-old ochre-processing workshop at Blombos cave, South Africa. *Science*, *334*(6053), 219-222. https://doi.org/10.1126/science.1211535

Herdson, O., Eerola, T., & Javadi, A.-H. (2023). Analysis and classification of music-induced states of sadness. *Emotion Review*, *15*(2), 99-117. https://doi.org/10.1177/17540739221140472

Hevner, K. (1935). Experimental studies of the affective value of colors and lines. *Journal of Applied Psychology*, *19*(4), 385-398. https://doi.org/10.1037/h0055538

Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, *48*(2), 246-268. https://doi.org/10.2307/1415746

Hevner, K. (1937). An experimental study of the affective value of sounds in poetry. *The American Journal of Psychology*, *49*(3), 419-434. https://doi.org/10.2307/1415776

Heyduk, R. G. (1975). Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception & Psychophysics*, *17*(1), 84-90. https://doi.org/10.3758/bf03204003

Hilgers, T. (2016). Aesthetic disinterestedness: Art, experience, and the self (1st ed.). Routledge. https://doi.org/10.4324/9781315696089

Hoemann, K., Nielson, C., Yuen, A., Gurera, J. W., Quigley, K. S., & Barrett, L. F. (2021). Expertise in emotion: A scoping review and unifying framework for individual differences in the mental representation of emotional experience. *Psychological Bulletin*, *147*(11), 1159-1183. https://doi.org/10.1037/bul0000327

Hogan, P. C. (2015). The idiosyncrasy of beauty: Aesthetic universals and the diversity of taste. In P. F. Bundgaard & F. Stjernfelt (Eds.), *Investigations into the phenomenology and the ontology of the work of art: What are artworks and how do we experience them?* (pp. 109-127). Springer. https://doi.org/10.1007/978-3-319-14090-2_7

Husselman, T.-A., Filho, E., Zugic, L. W., Threadgold, E., & Ball, L. J. (2024). Stimulus complexity can enhance art appreciation: Phenomenological and psychophysiological evidence for the pleasure-interest model of aesthetic liking. *Journal of Intelligence*, *12*(4), Article 42. https://www.mdpi.com/2079-3200/12/4/42

Ichikawa, S. (1985). Quantitative and structural factors in the judgment of pattern complexity. *Perception & Psychophysics*, *38*(2), 101-109. https://doi.org/10.3758/BF03198846

Iigaya, K., O'Doherty, J. P., & Starr, G. G. (2020). Progress and promise in neuroaesthetics. *Neuron*, *108*(4), 594-596. https://doi.org/10.1016/j.neuron.2020.10.022

Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., & O'Doherty, J. P. (2021). Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features. *Nature Human Behaviour*, *5*(6), 743-755. https://doi.org/10.1038/s41562-021-01124-6

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319-330. https://doi.org/10.1525/mp.2006.23.4.319

Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the psychology of aesthetics. *Leonardo*, *39*(2), 155-162. https://doi.org/10.1162/leon.2006.39.2.155

Jacobsen, T., & Beudt, S. (2017). Domain generality and domain specificity in aesthetic appreciation. *New Ideas in Psychology*, *47*, 97-102. https://doi.org/10.1016/j.newideapsych.2017.03.008

Jain, A., Schoeller, F., Horowitz, A., Hu, X., Yan, G., Salomon, R., & Maes, P. (2022). Aesthetic chills cause an emotional drift in valence and arousal. *Frontiers in Neuroscience*, *16*, Article 1013117. https://doi.org/10.3389/fnins.2022.1013117

Jakesch, M., & Leder, H. (2015). The qualitative side of complexity: Testing effects of ambiguity on complexity judgments. *Psychology of Aesthetics Creativity and the Arts*, *9*(3), 200-205. https://doi.org/10.1037/a0039350

Jakesch, M., Leder, H., & Forster, M. (2013). Image ambiguity and fluency. *PLOS One*, *8*(9), Article e74084. https://doi.org/10.1371/journal.pone.0074084

Jakubowski, K., Eerola, T., Blackwood Ximenes, A., Ma, W. K., Clayton, M., & Keller, P. E. (2020). Multimodal perception of interpersonal synchrony: Evidence from global and continuous ratings of improvised musical duo performances. *Psychomusicology: Music, Mind, and Brain*, *30*(4), 159-177. https://doi.org/10.1037/pmu0000264

Jankowski, T., Francuz, P., Oleś, P., & Chmielnicka-Kuter, E. (2020). The effect of temperament, expertise in art, and formal elements of paintings on their aesthetic appraisal. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(2), 209-223. https://doi.org/10.1037/aca0000211

Jankowski, T., Francuz, P., Oleś, P., Chmielnicka-Kuter, E., & Augustynowicz, P. (2020). The effect of painting beauty on eye movements. *Advances in Cognitive Psychology*, *16*(3), 213-227. https://doi.org/10.5709/acp-0298-4

Jensen, A., & Bonde, L. O. (2018). The use of arts interventions for mental health and wellbeing in health settings. *Perspectives in Public Health*, *138*(4), 209-214. https://doi.org/10.1177/1757913918772602

Jerram, M., Lee, A., Negreira, A., & Gansler, D. (2014). The neural correlates of the dominance dimension of emotion. *Psychiatry Research: Neuroimaging*, *221*(2), 135-141. https://doi.org/10.1016/j.pscychresns.2013.11.007

Johnson, O., & Knapp, R. H. (1963). Sex differences in aesthetic preferences. *The Journal of Social Psychology*, *61*(2), 279-301. https://doi.org/10.1080/00224545.1963.9919486

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235-266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., & Sloboda, J. A. (Eds.). (2010). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.

Kaimal, G., & Arslanbek, A. (2020). Indigenous and traditional visual artistic practices: Implications for art therapy clinical practice and research. *Frontiers in Psychology*, *11*, Article 1320. https://doi.org/10.3389/fpsyg.2020.01320

Kaisler, R. E., Marin, M. M., & Leder, H. (2020). Effects of emotional expressions, gaze, and head orientation on person perception in social situations. *Sage Open*, *10*(3), Article 2158244020940705. https://doi.org/10.1177/2158244020940705

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. Cambridge University Press.

Karim, A., Proulx, M. J., de Sousa, A. A., & Likova, L. T. (2022). Do we enjoy what we sense and perceive? A dissociation between aesthetic appreciation and basic perception of environmental objects or events. *Cognitive and Affective Behavioural Neuroscience*, *22*(5), 904-951. https://doi.org/10.3758/s13415-022-01004-0

Karjus, A., Canet Solà, M., Ohm, T., Ahnert, S. E., & Schich, M. (2023). Compression ensembles quantify aesthetic complexity and the evolution of visual art. *EPJ Data Science*, *12*(1), Article 21. https://doi.org/10.1140/epjds/s13688-023-00397-3

Keltner, D., & Oatley, K. (2022). Social functions of emotions in life and imaginative culture. *Evolutionary Studies in Imaginative Culture*, *6*(1), 1-20. https://doi.org/10.26613/esic.6.1.263

Khaw, M. W., & Freedberg, D. (2018). Continuous aesthetic judgment of image sequences. *Acta Psychologica*, *188*, 213-219. https://doi.org/10.1016/j.actpsy.2018.04.018

Kim, J., Shinkareva, S. V., & Wedell, D. H. (2017). Representations of modality-general valence for videos and music derived from fMRI data. *NeuroImage*, *148*, 42-54. https://doi.org/10.1016/j.neuroimage.2017.01.002

Kim, J., & Wedell, D. H. (2016). Comparison of physiological responses to affect eliciting pictures and music. *International Journal of Psychophysiology*, *101*, 9-17. https://doi.org/10.1016/j.ijpsycho.2015.12.011

Kiyokawa, H., & Hayashi, R. (2024). Commonalities and variations in emotion representation across modalities and brain regions. *Scientific Reports*, *14*(1), Article 20992. https://doi.org/10.1038/s41598-024-71690-y

Klein, S. B. (2014). What can recent replication failures tell us about the theoretical commitments of psychology? *Theory & Psychology*, *24*(3), 326-338. https://doi.org/10.1177/0959354314529616

Koelsch, S. (2011). Towards a neural basis of processing musical semantics. *Physics of Life Reviews*, *8*(2), 89-105. https://doi.org/10.1016/j.plrev.2011.04.004

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63-77. https://doi.org/10.1016/j.tics.2018.10.006

Konečni, V. J. (1996). Daniel E. Berlyne (1924-1976): Two decades later. *Empirical Studies of the Arts*, *14*(2), 129-142. https://doi.org/10.2190/DNFJ-8230-WBYX-JJB5

König, P., Wilming, N., Kietzmann, T. C., Ossandón, J. P., Onat, S., Ehinger, B. V., Gameiro, R. R., & Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movement Research*, *9*(5). https://doi.org/10.16910/jemr.9.5.3

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, *51*(13), 1457-1483. https://doi.org/10.1016/j.visres.2010.12.014

Kozbelt, A. (2022). The influence of expertise on aesthetics. In M. Nadal & O. Vartanian (Eds.), *The Oxford handbook of empirical aesthetics* (pp. 787-819). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.37

Kron, A., Pilkiw, M., Banaei, J., Goldstein, A., & Anderson, A. K. (2015). Are valence and arousal separable in emotional experience? *Emotion*, *15*(1), 35-44. https://doi.org/10.1037/a0038474

Kuijpers, M., Hakemulder, F., Bálint, K., Doicaru, M. & Tan, E. (2017). Towards a new understanding of absorbing reading experiences. In F. Hakemulder, M. Kuijpers, E. Tan, K. Bálint & M. Doicaru (Eds.), *Narrative Absorption* (pp. 29-47). John Benjamins Publishing Company. https://doi.org/10.1075/lal.27.03kui

Kuniecki, M., Woloszyn, K. B., Domagalik, A., & Pilarczyk, J. (2017). Effects of scene properties and emotional valence on brain activations: A fixation-related fMRI study. *Frontiers in Human Neuroscience*, *11*, Article 429. https://doi.org/10.3389/fnhum.2017.00429

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917-940. https://doi.org/10.1037/a0030811

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, *49*(2), 457-470. https://doi.org/10.3758/s13428-016-0715-3

Kyle-Davidson, C., Zhou, E. Y., Walther, D. B., Bors, A. G., & Evans, K. K. (2023). Characterising and dissecting human perception of scene complexity. *Cognition*, *231*, Article 105319. https://doi.org/10.1016/j.cognition.2022.105319

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 39-58.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (pp. A-8). Gainesville, FL: NIMH, Center for the Study of Emotion & Attention.

Latikka, R., Bergdahl, J., Savela, N., & Oksanen, A. (2023). AI as an artist? A two-wave survey study on attitudes toward using artificial intelligence in art. *Poetics*, *101*, Article 101839. https://doi.org/10.1016/j.poetic.2023.101839

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489-508. https://doi.org/10.1348/0007126042369811

Leder, H., Goller, J., Rigotti, T., & Forster, M. (2016). Private and shared taste in art and face appreciation. *Frontiers in Human Neuroscience*, *10*, Article 155. https://doi.org/10.3389/fnhum.2016.00155

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode – Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443-464. https://doi.org/10.1111/bjop.12084

Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. *Annual Review of Psychology*, *69*(1), 51-75. https://doi.org/10.1146/annurev-psych-122216-011740

Li, K., Weng, L., & Wang, X. (2021). The state of music therapy studies in the past 20 years: A bibliometric analysis. *Frontiers in Psychology*, *12*, Article 697726. https://doi.org/10.3389/fpsyg.2021.697726

Li, S., Zhang, P., & Wood, E. (2024). Familiarity and novelty in aesthetic appreciation: The case of intangible cultural heritage in China. *Annals of Tourism Research*, *105*, Article 103696. https://doi.org/10.1016/j.annals.2023.103696

Lill, K. (2012). Comparative theories of visual art and music: May I play you a picture? *McNair Scholars Journal*, *16*(1), 35-39.

Lin, H., Müller-Bardorff, M., Gathmann, B., Brieke, J., Mothes-Lasch, M., Bruchmann, M., Miltner, W. H. R., & Straube, T. (2020). Stimulus arousal drives amygdalar responses to emotional expressions across sensory modalities. *Scientific Reports*, *10*(1), Article 1898. https://doi.org/10.1038/s41598-020-58839-1

Liu, N., Tadesse, E., Zhang, M., Khalid, S., Li, W., & Gao, C. (2025). A review of event-related potential measuring aesthetic processing in visual aesthetic experience research. *Psychology of Aesthetics, Creativity, and the Arts*, Advance online publication. https://doi.org/10.1037/aca0000707

Locher, P. J. (2015). The aesthetic experience with visual art "at first glance" In P. F. Bundgaard & F. Stjernfelt (Eds.), *Investigations into the phenomenology and the ontology of the work of art: What are artworks and how do we experience them* (pp. 75–88). Springer Science + Business Media. https://doi.org/10.1007/978-3-319-14090-2_5

Locher, P. J., Overbeeke, K., & Wensveen, S. (2010) Aesthetic interaction: A framework. *Design Issues, 26*, 70-79. https://doi.org/10.1162/DESI_a_00017

Locher, P. J., Smith, J. K., & Smith, L. F. (2001). The influence of presentation format and viewer training in the visual arts on the perception of pictorial and aesthetic qualities of paintings. *Perception*, *30*(4), 449-465. https://doi.org/10.1068/p3008

Löw, A., Bradley, M. M., & Lang, P. J. (2013). Perceptual processing of natural scenes at rapid rates: Effects of complexity, content, and emotional arousal. *Cognitive Affective & Behavioral Neuroscience*, *13*(4), 860-868. https://doi.org/10.3758/s13415-013-0179-1

Luo, Y., Yu, Q., Wu, S., & Luo, Y.-j. (2025). Distinct neural bases of visual art- and music-induced aesthetic experiences. *NeuroImage*, *305*, Article 120962. https://doi.org/10.1016/j.neuroimage.2024.120962

Lyssenko, N., Redies, C., & Hayn-Leichsenring, G. U. (2016). Evaluating abstract art: Relation between term usage, subjective ratings, image properties and personality traits. *Frontiers in Psychology*, *7*, Article 973. https://doi.org/10.3389/fpsyg.2016.00973

Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta Psychologica*, *160*, 43-57. https://doi.org/10.1016/j.actpsy.2015.06.005

Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2018). Visual complexity and affect: Ratings reflect more than meets the eye. *Frontiers in Psychology*, *8*, Article 2368. https://doi.org/10.3389/fpsyg.2017.02368

Makin, A. D. J. (2017). The gap between aesthetic science and aesthetic experience. *Journal of Consciousness Studies*, *24*(1-2), 184-213.

Mandler, G. (2003). Emotion. In D. K. Freedheim (Ed.), *Handbook of psychology: History of psychology* (Vol. 1, pp. 157–175). John Wiley & Sons. https://doi.org/10.1002/0471264385.wei0108

Maness, E. B., Burk, J. A., McKenna, J. T., Schiffino, F. L., Strecker, R. E., & McCoy, J. G. (2022). Role of the locus coeruleus and basal forebrain in arousal and attention. *Brain Research Bulletin*, *188*, 47-58. https://doi.org/10.1016/j.brainresbull.2022.07.014

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*, 596-610. https://doi.org/10.3758/s13428-013-0379-1

Marin, M. M. (2015). Crossing boundaries: Toward a general model of neuroaesthetics. *Frontiers in Human Neuroscience*, *9*, Article 443. https://doi.org/10.3389/fnhum.2015.00443

Marin, M. M. (2022). The role of collative variables in aesthetic experiences. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (pp. 385-429). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.20

Marin, M. M., & Gingras, B. (2024). How music-induced emotions affect sexual attraction: Evolutionary implications. *Frontiers in Psychology*, *15*, Article 1269820. https://doi.org/10.3389/fpsyg.2024.1269820

Marin, M. M., Gingras, B., & Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion*, *12*(3), 618-631. https://doi.org/10.1037/a0025020

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience, 10*, Article 536. https://doi.org/10.3389/fnhum.2016.00536

Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLOS One*, *8*(8), Article e72412. https://doi.org/10.1371/journal.pone.0072412

Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica*, *163*, 38-58. https://doi.org/10.1016/j.actpsy.2015.10.002

Marin, M. M., & Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Personality and Individual Differences*, *125*, 80-90. https://doi.org/10.1016/j.paid.2017.12.027

Marin, M. M., & Leder, H. (2022). Gaze patterns reveal aesthetic distance while viewing art. *Annals of the New York Academy of Sciences, 1514*(1), 155-165. https://doi.org/10.1111/nyas.14792

Markovic, J., Anderson, A. K., & Todd, R. M. (2014). Tuning to the significant: Neural and genetic processes underlying affective enhancement of visual perception and memory. *Behavioural Brain Research*, *259*, 229-241. https://doi.org/10.1016/j.bbr.2013.11.018

Marković, S. (2011). Perceptual, semantic and affective dimensions of experience of abstract and representational paintings. *Psihologija*, *44*(3), 191-210. https://doi.org/10.2298/psi1103191m

Marković, S. (2012). Components of aesthetic experience: Aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *Iperception*, *3*(1), 1-17. https://doi.org/10.1068/i0450aap

Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *The Journal of Mind and Behavior*, *5*(1), 49-80. http://www.jstor.org/stable/43853015

Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 661-670. https://doi.org/10.1037/0096-1523.14.4.661

Martindale, C., & Moore, K. (1989). Relationship of musical preference to collative, ecological, and psychophysical variables. *Music Perception*, *6*(4), 431-445. https://doi.org/10.2307/40285441

Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *The American Journal of Psychology*, *103*(1), 53-80. https://doi.org/10.2307/1423259

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, *39*, Article e200. https://doi.org/10.1017/S0140525X15000667

Mauch, M., & Levy, M. (2011). Structural change on multiple time scales as a correlate of musical complexity. In A. Klapuri & C. Leider (Eds.), *Proceedings of the 12th International Conference on Music Information Retrieval ISMIR* (pp. 489-494). University of Miami.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209-237. https://doi.org/10.1080/02699930802204677

McCrae, R. R. (2024). A volitional account of aesthetic experience. *Frontiers in Psychology*, *15,* Article 148034. https://doi.org/10.3389/fpsyg.2024.1480304

McDermott, A. (2021). What was the first “art”? How would we know? *Proceedings of the National Academy of Sciences USA*, *118*(44), Article e2117561118. https://doi.org/10.1073/pnas.2117561118

McDonnell, M. (2007). Visual music. *Visual Music Marathon Program Catalog*, 1-19.

McMullen, P. T. (1974). Influence of number of different pitches and melodic redundancy on preference responses. *Journal of Research in Music Education*, *22*(3), 198-204. https://doi.org/10.2307/3345141

Mehl, K., Gugliano, M., & Belfi, A. M. (2023). The role of imagery and emotion in the aesthetic appeal of music, poetry, and paintings. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000623

Menninghaus, W., Schindler, I., Wagner, V., Wassiliwizky, E., Hanich, J., Jacobsen, T., & Koelsch, S. (2020). Aesthetic emotions are a key factor in aesthetic

evaluation: Reply to Skov and Nadal (2020). *Psychological Review*, *127*(4), 650-654. https://doi.org/10.1037/rev0000213

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*, *126*(2), 171-195. https://doi.org/10.1037/rev0000135

Merrill, J., Czepiel, A., Fink, L. T., Toelle, J., & Wald-Fuhrmann, M. (2021). The aesthetic experience of live concerts: Self-reports and psychophysiology. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(2), 134-151. https://doi.org/10.1037/aca0000390

Meyer, J., Thoma, G.-B., Kampschulte, L., & Köller, O. (2023). Openness to experience and museum visits: Intellectual curiosity, aesthetic sensitivity, and creative imagination predict the frequency of visits to different types of museums. *Journal of Research in Personality*, *103*, Article 104352. https://doi.org/10.1016/j.jrp.2023.104352

Miller, C. A., & Hübner, R. (2020). Two routes to aesthetic preference, one route to aesthetic inference. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(2), 237-249. https://doi.org/10.1037/aca0000241

Miller, C. A., & Hübner, R. (2023). The relations of empathy and gender to aesthetic response and aesthetic inference of visual artworks. *Empirical Studies of the Arts*, *41*(1), 188-215. https://doi.org/10.1177/02762374221095701

Miller, G. F. (2001). Aesthetic fitness: How sexual selection shaped artistic virtuosity as a fitness indicator and aesthetic preferences as mate choice criteria. *Bulletin of Psychology and the Arts*, *2*(1), 20-25.

Mittelberg, I. (2013): Balancing acts: Image schemas and force dynamics as experiential essence in pictures by Paul Klee and their gestural enactments, In B. Dancygier, M. Bokrent & J. Hinnell (Eds.), *Language and the Creative Mind* (pp. 325-346). Stanford: Center for the Study of Language and Information.

Miu, A. C., Pițur, S., & Szentágotai-Tătar, A. (2016). Aesthetic emotions across arts: A comparison between painting and music. *Frontiers in Psychology*, *6*, Article 1951. https://doi.org/10.3389/fpsyg.2015.01951

Moles, A. A. (1958). *Theorie de linformation et perception esthetique*. Flammarion.

Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, *23*(3), 257-278. https://doi.org/10.1521/soco.2005.23.3.257

Moore, K. M., & West, A. N. (2012). Global perception, meaning, and aesthetic appreciation. *Empirical Studies of the Arts*, *30*(1), 23-38. https://doi.org/10.2190/EM.30.1.d

Munoz, N. E., & Blumstein, D. T. (2012). Multisensory perception in uncertain environments. *Behavioral Ecology*, *23*(3), 457-462. https://doi.org/10.1093/beheco/arr220

Muth, C., Raab, M. H., & Carbon, C. C. (2015). The stream of experience when watching artistic movies. Dynamic aesthetic effects revealed by the Continuous Evaluation Procedure (CEP). *Frontiers in Psychology*, *6*, Article 365. https://doi.org/10.3389/fpsyg.2015.00365

Myszkowski, N. (2022). Aesthetic sensitivity. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (pp. 834-852). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.40

Nadal, M., & Chatterjee, A. (2019). Neuroaesthetics and art's diversity and universality. *WIREs Cognitive Science*, *10*(3), Article e1487. https://doi.org/10.1002/wcs.1487

Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. J. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts*, *28*(2), 173-191. https://doi.org/10.2190/EM.28.2.d

Nadal, M., & Skov, M. (2025). The sensory valuation account of aesthetic experience. *Nature Reviews Psychology*, *4*(1), 49-63. https://doi.org/10.1038/s44159-024-00385-y

Nadal, M., & Urena, E. (2022). One hundred years of empirical aesthetics: Fechner to Berlyne (1876-1976). In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (pp. 39-82). Oxford University Press.

Nagle, F., & Lavie, N. (2020). Predicting human complexity perception of real-world scenes. *Royal Society Open Science*, *7*(5), Article 191487. https://doi.org/10.1098/rsos.191487

Nater, U. M., Abbruzzese, E., Krebs, M., & Ehlert, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology*, *62*(2), 300-308. https://doi.org/10.1016/j.ijpsycho.2006.05.011

Nath, S. S., Brändle, F., Schulz, E., Dayan, P., & Brielmann, A. (2024). Relating objective complexity, subjective complexity, and beauty in binary pixel patterns. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000657

Nava, E., Giraud, M., & Bolognini, N. (2024). The emergence of the multisensory brain: From the womb to the first steps. *iScience*, *27*(1), Article 108758. https://doi.org/10.1016/j.isci.2023.108758

Nicki, R. M., & Moss, V. (1975). Preference for non-representational art as a function of various measures of complexity. *Canadian Journal of Psychology / Revue canadienne de psychologie*, *29*(3), 237-249. https://doi.org/10.1037/h0082029

Nineuil, C., Dellacherie, D., & Samson, S. (2022). French adaptation of a film music stimulus set: Normative emotional ratings of valence and arousal prompted by music excerpts. *Psychology of Music*, *50*(5), 1707-1715, Article 03057356211050683. https://doi.org/10.1177/03057356211050683

Niu, Y., Todd, R., & Anderson, A. K. (2012). Affective salience can reverse the effects of stimulus-driven salience on eye movements in complex scenes. *Frontiers in Psychology*, *3*, Article 336. https://doi.org/10.3389/fpsyg.2012.00336

Nordström, H., & Laukka, P. (2019). The time course of emotion recognition in speech and music. *The Journal of the Acoustical Society of America*, *145*(5), 3058-3074. https://doi.org/10.1121/1.5108601

North, A., & Hargreaves, D. (2008). *The social and applied psychology of music*. Oxford University Press.

North, A. C., & Hargreaves, D. J. (2000). Collative variables versus prototypicality. *Empirical Studies of the Arts*, *18*(1), 13-17. https://doi.org/10.2190/k96d-085m-t07y-61ab

Ogg, M., Sears, D. R. W., Marin, M. M., & McAdams, S. (2017). Psychophysiological indices of music-evoked emotions in musicians. *Music Perception*, *35*(1), 38-59. https://doi.org/10.1525/mp.2017.35.1.38

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247-265. https://doi.org/10.1016/j.biopsycho.2007.11.006

Orr, M. G., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception*, *22*(4), 583-611. https://doi.org/10.1525/mp.2005.22.4.583

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois press.

Palczak, J., Blumenthal, D., & Delarue, J. (2020). Influence of sensory complexity on preferences for novel gourmet dairy desserts. Does Berlyne's theory apply to desserts? *Food Quality and Preference*, *84*, Article 103957. https://doi.org/10.1016/j.foodqual.2020.103957

Pallas, S. L., & Mao, Y.-T. (2012). The evolution of multisensory neocortex. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 627-641). MIT Press.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*(1), 77-107. https://doi.org/10.1146/annurev-psych-120710-100504

Papenmeier, F., Dagit, G., Wagner, C., & Schwan, S. (2024). Is it art? Effects of framing images as art versus non-art on gaze behavior and aesthetic judgments. *Psychology of Aesthetics, Creativity, and the Arts*, *18*(4), 642-653. https://doi.org/10.1037/aca0000466

Park, C., & Kim, J. (2024). Taste the music: Modality-general representation of affective states derived from auditory and gustatory stimuli. *Cognition*, *249*, Article 105830. https://doi.org/10.1016/j.cognition.2024.105830

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, *11*(2), 265-279. https://doi.org/10.1177/1745691615621274

Pease, A., Mahmoodi, K., & West, B. J. (2018). Complexity measures of music. *Chaos, Solitons & Fractals*, *108*, 82-86. https://doi.org/10.1016/j.chaos.2018.01.021

Pelowski, M., & Akiba, F. (2011). A model of art perception, evaluation and emotion in transformative aesthetic experience. *New Ideas in Psychology*, *29*(2), 80-97. https://doi.org/10.1016/j.newideapsych.2010.04.001

Pelowski, M., Cabbai, G., Brinkmann, H., Mikuni, J., Hegelmaier, L. M., Forster, M., Rosenberg, R., & Leder, H. (2020). The kitsch switch—or (when) do experts dislike Thomas Kinkade art? A study of time-based evaluation changes in top-down versus bottom-up assessment. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000302

Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with 'real' and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(3), 245-264. https://doi.org/10.1037/aca0000141

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me… delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, *21*, 80-125. https://doi.org/10.1016/j.plrev.2017.02.003

Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in Human Neuroscience*, *10*, Article 160. https://doi.org/10.3389/fnhum.2016.00160

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148-158. https://doi.org/10.1038/nrn2317

Pessoa, L. (2022). *The entangled brain: How perception, cognition, and emotion are woven together*. The MIT Press. https://doi.org/10.7551/mitpress/14636.001.0001

Petrolini, V., & Viola, M. (2020). Core affect dynamics: Arousal as a modulator of valence. *Review of Philosophy and Psychology*, *11*(4), 783-801. https://doi.org/10.1007/s13164-020-00474-w

Pihko, E., Virtanen, A., Saarinen, V. M., Pannasch, S., Hirvenkari, L., Tossavainen, T., Haapala, A., & Hari, R. (2011). Experiencing art: The influence of expertise and painting abstraction level. *Frontiers in Human Neuroscience*, *5*, Article 94. https://doi.org/10.3389/fnhum.2011.00094

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception*, *30*(1), 71-83. https://doi.org/10.1525/mp.2012.30.1.71

Plazak, J., & Huron, D. (2011). The first three seconds: Listener knowledge gained from brief musical excerpts. *Musicae Scientiae*, *15*(1), 29-44. https://doi.org/10.1177/1029864910391455

Pombo, M., & Pelli, D. G. (2022). Aesthetics: It's beautiful to me. *Current Biology*, *32*(8), R378-R379. https://doi.org/10.1016/j.cub.2022.03.002

Potter, R. F., & Choi, J. (2006). The effects of auditory structural complexity on attitudes, attention, arousal, and memory. *Media Psychology*, *8*(4), 395-419. https://doi.org/10.1207/s1532785xmep0804_4

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*(3), 492-512. https://doi.org/10.1016/j.biopsycho.2012.02.007

Quadt, L., Critchley, H., & Nagai, Y. (2022). Cognition, emotion, and the central autonomic network. *Autonomic neuroscience : Basic & Clinical*, *238*, 102948. https://doi.org/10.1016/j.autneu.2022.102948

Pressing, J. (1999). Cognitive complexity and the structure of musical patterns. In *Proceedings of the 4th Conference of the Australasian Cognitive Science Society*, Newcastle, Australia. http://dub.ucsd.edu/Mu206/CogComplex-music.pdf

Ragot, M., Martin, N., & Cojean, S. (2020). AI-generated vs. human artworks. A perception bias towards artificial intelligence? In R. Bernhaupt, F. Mueller, et al. (Eds.), *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1-10), Association for Computing Machinery. https://doi.org/10.1145/3334480.3382892

Ravignani, A. (2018). Darwin, sexual selection, and the origins of music. *Trends in Ecology & Evolution*, *33*(10), 716-719. https://doi.org/10.1016/j.tree.2018.07.006

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364-382. https://doi.org/10.1207/s15327957pspr0804_3

Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Frontiers in Human Neuroscience*, *9*, Article 218. https://doi.org/10.3389/fnhum.2015.00218

Redies, C., Amirshahi, S.A., Koch, M., & Denzler, J. (2012). PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In A. Fusiello, V. Murino, & R. Cucchiara (Eds.) *Computer Vision – ECCV 2012. Workshops and demonstrations. ECCV 2012. Lecture notes in*

computer science (Vol. 7583). Springer. https://doi.org/10.1007/978-3-642-33863-2_54

Redies, C., Bartho, R., Koßmann, L. *et al.* (2025). A toolbox for calculating quantitative image properties in aesthetics research. *Behavior Research*, *57*, Article 117. https://doi.org/10.3758/s13428-025-02632-3

Redies, C., Grebenkina, M., Mohseni, M., Kaduhm, A., & Dobel, C. (2020). Global image properties predict ratings of affective pictures. *Frontiers in Psychology*, *11*. Article 953. https://doi.org/10.3389/fpsyg.2020.00953

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, *67*(3), 525-539. https://doi.org/10.1037/0022-3514.67.3.525

Rind, M. (2002). The concept of disinterestedness in eighteenth-century British aesthetics. *Journal of the History of Philosophy*, *40*(1), 67-87.

Robles, K. E., Roberts, M., Viengkham, C., Smith, J. H., Rowland, C., Moslehi, S., Stadlober, S., Lesjak, A., Lesjak, M., Taylor, R. P., Spehar, B., & Sereno, M. E. (2021). Aesthetics and psychological effects of fractal based design. *Frontiers in Psychology*, *12*, Article 699962. https://doi.org/10.3389/fpsyg.2021.699962

Rodriguez, R. M., Fekete, A., Silvia, P. J., & Cotter, K. N. (2024). The art of feeling different: Exploring the diversity of emotions experienced during an art museum visit. *Psychology of Aesthetics, Creativity, and the Arts*, *18*(3), 303-314. https://doi.org/10.1037/aca0000443

Rolls, E. T. (2017). Neurobiological foundations of aesthetics and art. *New Ideas in Psychology*, *47*, 121-135. https://doi.org/10.1016/j.newideapsych.2017.03.005

Rösler, L., End, A., & Gamer, M. (2017). Orienting towards social features in naturalistic scenes is reflexive. *PLOS One*, *12*(7), Article e0182037. https://doi.org/10.1371/journal.pone.0182037

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217-236. https://doi.org/10.1068/p5117

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-1178. https://doi.org/10.1037/h0077714

Russell J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145-172. https://doi.org/10.1037/0033-295x.110.1.145

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805-819. https://doi.org/10.1037/0022-3514.76.5.805

Russell, P. A., & George, D. A. (1990). Relationships between aesthetic response scales applied to paintings. *Empirical Studies of the Arts*, *8*(1), 15-30. https://doi.org/10.2190/au1r-6uxe-t14r-04wq

Saklofske, D. H. (1975). Visual aesthetic complexity, attractiveness and diversive exploration. *Perceptual and Motor Skills*, *41*(3), 813-814. https://doi.org/10.2466/pms.1975.41.3.813

Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: How musical sounds become rewarding. *Trends in Cognitive Sciences*, *19*(2), 86-91. https://doi.org/10.1016/j.tics.2014.12.001

Sander, D. (2013). Models of emotion: The affective neuroscience approach. In J. Armony & P. Vuilleumier (Eds.), *The Cambridge handbook of human affective*

neuroscience (pp. 5–53). Cambridge University Press. https://doi.org/10.1017/CBO9780511843716.003

Sano, T., Gondo, M., Osaka, Y., & Kawabata, H. (2025). The effects of image resolution and exposure duration on facial beauty and ugliness evaluations. *Acta Psychologica*, *252*, Article 104635. https://doi.org/10.1016/j.actpsy.2024.104635

Sanocki, T., Nguyen, T., Shultz, S., & Defant, J. (2023). Novel scene understanding, from gist to elaboration. *Visual Cognition*, *31*(3), 188-215. https://doi.org/10.1080/13506285.2023.2221047

Sara S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews. Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Sarasso, P., Neppi-Modona, M., Sacco, K., & Ronga, I. (2020). “Stopping for knowledge”: The sense of beauty in the perception-action cycle. *Neuroscience & Biobehavioral Reviews*, *118*, 723-738. https://doi.org/10.1016/j.neubiorev.2020.09.004

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*, Article 511. https://doi.org/10.3389/fpsyg.2013.00511

Schepman, A., & Rodway, P. (2021). Concreteness of semantic interpretations of abstract and representational artworks. *Acta Psychologica*, *215*, Article 103269. https://doi.org/10.1016/j.actpsy.2021.103269

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*, *33*(3), 239-251. https://doi.org/10.1080/0929821042000317822

Scherer, K. R., Trznadel, S., Fantini, B., & Coutinho, E. (2019). Assessing emotional experiences of opera spectators in situ. *Psychology of Aesthetics Creativity and the Arts*, *13*(3), 244-258. https://doi.org/10.1037/aca0000163

Schimmack, U. (2005). Attentional interference effects of emotional pictures: Threat, negativity, or arousal? *Emotion*, *5*(1), 55-66. https://doi.org/10.1037/1528-3542.5.1.55

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS One*, *12*(6), Article e0178899. https://doi.org/10.1371/journal.pone.0178899

Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, *51*(3), 154-165. https://doi.org/10.1080/00049539908255353

Schubert, E. (2013). Emotion felt by the listener and expressed by the music: Literature review and theoretical perspectives. *Frontiers in Psychology*, *4*, Article 837. https://doi.org/10.3389/fpsyg.2013.00837

Schubert, E. (2024). The aesthetic emotion lexicon: A literature review of emotion words used by researchers to describe aesthetic experiences. *Empirical Studies of the Arts*, *42*(1), 3-37. https://doi.org/10.1177/02762374221143728

Schueller, H. M. (1953). Correspondences between music and the sister arts, according to 18th century aesthetic theory. *The Journal of Aesthetics and Art Criticism*, *11*(4), 334-359. https://doi.org/10.2307/426456

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1), 7-13. https://doi.org/10.1111/1467-9280.01411

Schwabe, K., Menzel, C., Mullin, C., Wagemans, J., & Redies, C. (2018). Gist perception of image composition in abstract artworks. *i-Perception*, *9*(3), Article 2041669518780797. https://doi.org/10.1177/2041669518780797

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience*, *23*(7), 439-452. https://doi.org/10.1038/s41583-022-00587-4

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews*, *7*(3), 269-284. https://doi.org/10.1016/j.plrev.2010.04.006

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487-510. https://doi.org/10.1146/annurev-psych-122216-011845

Sidhu, D. M., McDougall, K. H., Jalava, S. T., & Bodner, G. E. (2018). Prediction of beauty and liking ratings for abstract and representational paintings using subjective and objective measures. *PLOS One*, *13*(7), Article e0200431. https://doi.org/10.1371/journal.pone.0200431

Silvia, P. J. (2005). Emotional responses to art: From collation and arousal to cognition and emotion. *Review of General Psychology*, *9*(4), 342-357. https://doi.org/10.1037/1089-2680.9.4.342

Silvia, P. J. (2006). *Exploring the psychology of interest*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195158557.001.0001

Silvia, P. J. (2009). Looking past pleasure: Anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 48-51. https://doi.org/10.1037/a0014632

Šimić, G., Tkalčić, M., Vukić, V., Mulc, D., Španić, E., Šagud, M., Olucha-Bordonau, F. E., Vukšić, M., & R. Hof, P. (2021). Understanding emotions: Origins and roles of the amygdala. *Biomolecules*, *11*(6), Article 823. https://www.mdpi.com/2218-273X/11/6/823

Simola, J., Le Fevre, K., Torniainen, J., & Baccino, T. (2015). Affective processing in natural scene viewing: Valence and arousal interactions in eye-fixation-related potentials. *NeuroImage*, *106*, 21-33. https://doi.org/10.1016/j.neuroimage.2014.11.030

Skov, M., & Nadal, M. (2020). There are no aesthetic emotions: Comment on Menninghaus et al. (2019). *Psychological Review*, *127*(4), 640-649. https://doi.org/10.1037/rev0000187

Skov, M., & Nadal, M. (2021a). The nature of beauty: Behavior, cognition, and neurobiology. *Annals of the New York Academy of Sciences*, *1488*(1), 44-55. https://doi.org/10.1111/nyas.14524

Skov, M., & Nadal, M. (2021b). The nature of perception and emotion in aesthetic appreciation: A response to Makin's challenge to empirical aesthetics. *Psychology of Aesthetics Creativity and the Arts*, *15*(3), 470-483. https://doi.org/10.1037/aca0000278

Skripkauskaite, S., Mihai, I., & Koldewyn, K. (2023). Attentional bias towards social interactions during viewing of naturalistic scenes. *Quarterly Journal of Experimental Psychology*, *76*(10), 2303-2311. https://doi.org/10.1177/17470218221140879

Smith, J. K., & Smith, L. F. (2001). Spending time on art. *Empirical Studies of the Arts*, *19*(2), 229-236. https://doi.org/10.2190/5mqm-59jh-x21r-jn5j

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174-215. https://doi.org/10.1037/0278-7393.6.2.174

Song, J., Kwak, Y., & Kim, C.-Y. (2021). Familiarity and novelty in aesthetic preference: The effects of the properties of the artwork and the beholder. *Frontiers in Psychology*, *12*, Article 694927. https://doi.org/10.3389/fpsyg.2021.694927

Sorek, Y., Nahari, T., & Pertzov, Y. (2024). Artistic expertise and free viewing of modern art. *Art & Perception*, *12*(4), 346-379. https://doi.org/10.1163/22134913-bja10062

Specker, E., Fried, E. I., Rosenberg, R., & Leder, H. (2021). Associating with art: A network model of aesthetic effects. *Collabra: Psychology*, *7*(1), Article 24085. https://doi.org/10.1525/collabra.24085

Spee, B. T. M., Arato, J., Mikuni, J., Tran, U. S., Pelowski, M., & Leder, H. (2024). The dynamics of experiencing Gestalt and Aha in cubist art: Pupil responses and art evaluations show a complex interplay of task, stimuli content, and time course. *Frontiers in Psychology*, *15*, Article 1192565. https://doi.org/10.3389/fpsyg.2024.1192565

Stamkou, E., Keltner, D., Corona, R., Aksoy, E., & Cowen, A. S. (2024). Emotional palette: A computational mapping of aesthetic experiences evoked by visual art. *Scientific Reports*, *14*(1), Article 19932. https://doi.org/10.1038/s41598-024-69686-9

Stechow, W. (1953). Problems of structure in some relations between the visual arts and music. *The Journal of Aesthetics and Art Criticism*, *11*(4), 324-333. https://doi.org/10.2307/426455

Stenson, H. H. (1966). The physical factor structure of random forms and their judged complexity. *Perception & Psychophysics*, *1*(9), 303-310. https://doi.org/10.3758/BF03215793

Storbeck, J., & Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass*, *2*(5), 1824-1843. https://doi.org/10.1111/j.1751-9004.2008.00138.x

Straffon, L. M. (2019). Evolution and the origins of visual art: An archaeological perspective. In A. M. Prentiss (Ed.), *Handbook of Evolutionary Research in Archaeology* (pp. 407-435). Springer International Publishing. https://doi.org/10.1007/978-3-030-11117-5_20

Street, N., Forsythe, A. M., Reilly, R., Taylor, R., & Helmy, M. S. (2016). A complex story: Universal preference vs. individual differences shaping aesthetic response to fractals patterns. *Frontiers in Human Neuroscience*, *10*, Article 213. https://doi.org/10.3389/fnhum.2016.00213

Streich, S. (2007). *Music complexity: A multi-faceted description of audio content.* [Unpublished doctoral dissertation]. University of Pompeu Fabra.

Strijbosch, W., Vessel, E. A., Welke, D., Mitas, O., Gelissen, J., & Bastiaansen, M. (2022). On the neuronal dynamics of aesthetic experience: Evidence from electroencephalographic oscillatory dynamics. *Journal of Cognitive Neuroscience*, *34*(3), 461-479. https://doi.org/10.1162/jocn_a_01812

Sun, L., Yamasaki, T., Aizawa, K. (2015). Relationship between visual complexity and aesthetics: Application to beauty prediction of photos. In L. Agapito, M. Bronstein, C. Rother (Eds.) Computer Vision - ECCV 2014 Workshops. Lecture Notes in Computer Science (Vol. 8925), Springer. https://doi.org/10.1007/978-3-319-16178-5_2

Sun, Z., & Firestone, C. (2021). Curious objects: How visual complexity guides attention and engagement. *Cognitive Science*, *45*(4), Article e12933. https://doi.org/10.1111/cogs.12933

Sutton, T. M., Herbert, A. M., & Clark, D. Q. (2019). Valence, arousal, and dominance ratings for facial stimuli. *Quarterly Journal of Experimental Psychology*, *72*(8), 2046-2055. https://doi.org/10.1177/1747021819829012

Swami, V., & Furnham, A. (2022). The influence of personality on aesthetic preferences. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (pp. 820-833). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.36

T. Zaatar, M., Alhakim, K., Enayeh, M., & Tamer, R. (2024). The transformative power of music: Insights into neuroplasticity, health, and disease. *Brain, Behavior, & Immunity - Health*, *35*, Article 100716. https://doi.org/10.1016/j.bbih.2023.100716

Takano, R., & Nomura, M. (2022). Neural representations of awe: Distinguishing common and distinct neural mechanisms. *Emotion*, *22*(4), 669-677. https://doi.org/10.1037/emo0000771

Tan, E. S. (2000). Emotion, art and the humanities. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (pp. 116-136). Guilford Press.

Tan, S.-L., Cohen, A. J., Lipscomb, S. D., & Kendall, R. A. (Eds.). (2013). *The Psychology of Music in Multimedia*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608157.001.0001.

Taylor, R. P., Spehar, B., Wise, J. A., Clifford, C. W., Newell, B. R., Hagerhall, C. M., Purcell, T., & Martin, T. P. (2005). Perceptual and physiological responses to the visual complexity of fractal patterns. *Nonlinear Dynamics, Psychology, and Life Sciences*, *9*(1), 89-114.

Tetzlaff, B.-O., Weiler, S. M., Herzberg, P. Y., & Jacobsen, T. (2024). Assessing the desire for aesthetics: Adaptation and validation of the desire for aesthetics scale in German (DFAS-G). *Acta Psychologica*, *247*, Article 104322. https://doi.org/10.1016/j.actpsy.2024.104322

Texier, P.-J., Porraz, G., Parkington, J., Rigaud, J.-P., Poggenpoel, C., Miller, C., Tribolo, C., Cartwright, C., Coudenneau, A., Klein, R., Steele, T., & Verna, C. (2010). A Howiesons Poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at Diepkloof Rock Shelter, South Africa. *Proceedings of the National Academy of Sciences USA*, *107*(14), 6180-6185. https://doi.org/10.1073/pnas.0913047107

Tiihonen, M., Brattico, E., Maksimainen, J., Wikgren, J., & Saarikallio, S. (2017). Constituents of music and visual-art related pleasure-a critical integrative literature review. *Frontiers in Psychology*, *8*, Article 1218. https://doi.org/10.3389/fpsyg.2017.01218

Tiihonen, M., Jacobsen, T., Trusbak Haumann, N., Saarikallio, S., & Brattico, E. (2022). I know what i like when i see it: Likability is distinct from pleasantness since early stages of multimodal emotion evaluation. *PLOS One*, *17*(9), Article e0274556. https://doi.org/10.1371/journal.pone.0274556

Tinio, P. P. L., & Gartus, A. (2018). Characterizing the emotional response to art beyond pleasure: Correspondence between the emotional characteristics of artworks and viewers' emotional responses. In J. F. Christensen & A. Gomila (Eds.), *Progress in Brain Research* (Vol. 237, pp. 319-342). Elsevier. https://doi.org/10.1016/bs.pbr.2018.03.005

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 114-137). The Guilford Press.

Trupp, M. D., Bignardi, G., Specker, E., Vessel, E. A., & Pelowski, M. (2023). Who benefits from online art viewing, and how: The role of pleasure, meaningfulness, and trait aesthetic responsiveness in computer-based art interventions for well-being. *Computers in Human Behavior*, *145*, Article 107764. https://doi.org/10.1016/j.chb.2023.107764

Tuch, A. N., Bargas-Avila, J. A., & Opwis, K. (2010). Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, *26*(6), 1831-1837. https://doi.org/10.1016/j.chb.2010.07.016

Van de Cruys, S., Frascaroli, J., & Friston, K. (2024). Order and change in art: Towards an active inference account of aesthetic experience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1895), Article 20220411. https://doi.org/10.1098/rstb.2022.0411

Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *i-Perception*, *2*(9), 1035-1062. https://doi.org/10.1068/i0466aap

van Den Bosch, I., Salimpoor, V., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience*, *7*, Article 534. https://doi.org/10.3389/fnhum.2013.00534

van Der Lee, C., & Van Enschot, R. (2022). Understanding emotional responses to visual aesthetic artefacts: The SECMEA mechanisms. *Visual Communication*, Article 14703572221090536. https://doi.org/10.1177/14703572221090536

van Geert, E., & Wagemans, J. (2020). Order, complexity, and aesthetic appreciation. *Psychology of Aesthetics Creativity and the Arts*, *14*(2), 135-154. https://doi.org/10.1037/aca0000224

van Geert, E., & Wagemans, J. (2021). Order, complexity, and aesthetic preferences for neatly organized compositions. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(3), 484-504. https://doi.org/10.1037/aca0000276

van Leeuwen, J. E. P., Boomgaard, J., Bzdok, D., Crutch, S. J., & Warren, J. D. (2022). More than meets the eye: Art engages the social brain. *Frontiers in Neuroscience*, *16*, Article 738865. https://doi.org/10.3389/fnins.2022.738865

van Paasschen, J., Bacci, F., & Melcher, D. P. (2015). The influence of art expertise and training on emotion and preference ratings for representational and abstract artworks. *PLOS One*, *10*(8), Article e0134241. https://doi.org/10.1371/journal.pone.0134241

van Wassenhove, V., Ghazanfar, A. A., Munhall, K. G., & Schröder, C. E. (2012). Bridging the gap between human and nonhuman studies of audiovisual integration. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 153-167). MIT Press.

Vara Sánchez, C. (2022). Enacting the aesthetic: A model for raw cognitive dynamics. *Phenomenology and the Cognitive Sciences*, *21*(2), 317-339. https://doi.org/10.1007/s11097-021-09737-y

Varella, M. A. C., de Souza, A. A. L., & Ferreira, J. H. B. P. (2011). Evolutionary aesthetics and sexual selection in the evolution of rock art aesthetics. *Rock Art Research*, *28*(2), 153-186. https://search.informit.org/doi/10.3316/ielapa.495427368208166

Verhaeghen, P. (2018). Once more, with feeling: The role of familiarity in the aesthetic response. *The Psychological Record*, *68*(3), 379-384. https://doi.org/10.1007/s40732-018-0312-1

Verhavert, S., Wagemans, J., & Augustin, M. D. (2018). Beauty in the blink of an eye: The time course of aesthetic experiences. *British Journal of Psychology*, *109*(1), 63-84. https://doi.org/10.1111/bjop.12258

Vessel, E. A. (2022). Neuroaesthetics. In S. Della Sala (Ed.), *Encyclopedia of Behavioral Neuroscience, 2nd edition* (pp. 661-670). Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.24104-7

Vessel, E. A., Ishizu, T., & Bignardi, G. (2023). Neural correlates of visual aesthetic appeal. In M. Skov & M. Nadal (Eds.), *The Routledge international handbook of neuroaesthetics* (pp. 103–133). Routledge.

Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, *179*, 121-131. https://doi.org/10.1016/j.cognition.2018.06.009

Vessel, E. A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *Journal of Vision*, *10*(2), 11-14. https://doi.org/10.1167/10.2.18

Vessel, E. A., Starr, G. G., & Rubin, N. (2012). The brain on art: Intense aesthetic experience activates the default mode network. *Frontiers in Human Neuroscience*, *6*, 1-17. https://doi.org/10.3389/fnhum.2012.00066

Villani, D., Morganti, F., Cipresso, P., Ruggi, S., Riva, G., & Gilli, G. (2015). Visual exploration patterns of human figures in action: An eye tracker study with art paintings. *Frontiers in Psychology*, *6*, Article 1636. https://doi.org/10.3389/fpsyg.2015.01636

Vissers, N., Moors, P., Genin, D., & Wagemans, J. (2020). Exploring the role of complexity, content and individual differences in aesthetic reactions to semi-abstract art photographs. *Art & Perception*, *8*(1), 89-119. https://doi.org/10.1163/22134913-20191139

Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology*, *71*(1), 74-79. https://doi.org/10.1037/h0022619

Wagner, V., Menninghaus, W., Hanich, J., & Jacobsen, T. (2014). Art schema effects on affective experience: The case of disgusting images. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(2), 120-129. https://doi.org/10.1037/a0036126

Wagner, V., Scharinger, M., Knoop, C. A., & Menninghaus, W. (2021). Effects of continuous self-reporting on aesthetic evaluation and emotional responses. *Poetics*, *85*, Article 101497. https://doi.org/10.1016/j.poetic.2020.101497

Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music*. MIT Press.

Watt, R. J., & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae Scientiae*, *2*(1), 33-53. https://doi.org/10.1177/102986499800200103

Weigand, R., & Jacobsen, T. (2023). Looking at life through rose-colored glasses: Dispositional positive affect is related to the intensity of aesthetic experiences. *The Journal of Positive Psychology*, *18*(4), 517-530. https://doi.org/10.1080/17439760.2022.2070530

Weiß, C., & Müller, M. (2014). Quantifying and visualizing tonal complexity. In T. Klouche, & E. R. Miranda (Eds.), *Proceedings of the Conference on Interdisciplinary Musicology (CIM14)* (pp. 184-188), Staatliches Institut für Musikforschung Berlin.

Welke, D., & Vessel, E. A. (2022). Naturalistic viewing conditions can increase task engagement and aesthetic preference but have only minimal impact on EEG quality. *NeuroImage*, *256*, Article 119218. https://doi.org/10.1016/j.neuroimage.2022.119218

Welter, M., & Lotte, F. (2024). Ecological decoding of visual aesthetic preference with oscillatory electroencephalogram features - A mini-review. *Frontiers in Neuroergonomics*, *5,* Article 1341790. https://doi.org/10.3389/fnrgo.2024.1341790

Wilkinson, Z., Cunningham, R., & Elliott, M. A. (2024). The influence of empathy on the perceptual response to visual art. *Psychology of Aesthetics, Creativity, and the Arts, 18*(3), 259–268. https://doi.org/10.1037/aca0000418

Wu, D. W.-L., Anderson, N. C., Bischof, W. F., & Kingstone, A. (2014). Temporal dynamics of eye movements are related to differences in scene complexity and clutter. *Journal of Vision*, *14*(9), 8-8. https://doi.org/10.1167/14.9.8

Yanagisawa, H., Kawamata, O., & Ueda, K. (2019). Modeling emotions associated with novelty at variable uncertainty levels: A Bayesian approach. *Frontiers in Computational Neuroscience*, *13*, Article 2. https://doi.org/10.3389/fncom.2019.00002

Yang, W., Makita, K., Nakao, T., Kanayama, N., Machizawa, M. G., Sasaoka, T., Sugata, A., Kobayashi, R., Hiramoto, R., Yamawaki, S., Iwanaga, M., & Miyatani, M. (2018). Affective auditory stimulus database: An expanded version of the International Affective Digitized Sounds (IADS-E). *Behavior Research Methods*, *50*(4), 1415-1429. https://doi.org/10.3758/s13428-018-1027-6

Yik, M., Mues, C., Sze, I. N. L., Kuppens, P., Tuerlinckx, F., De Roover, K., Kwok, F. H. C., Schwartz, S. H., Abu-Hilal, M., Adebayo, D. F., Aguilar, P., Al-Bahrani, M., Anderson, M. H., Andrade, L., Bratko, D., Bushina, E., Choi, J. W., Cieciuch, J., Dru, V., Evers, U., … Russell, J. A. (2023). On the relationship between valence and arousal in samples across the globe. *Emotion*, *23*(2), 332–344. https://doi.org/10.1037/emo0001095

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2, Pt.2), 1-27. https://doi.org/10.1037/h0025848

Zhou, E. Y., Wilder, J., Damiano, C., & Walther, D. B. (2023). Neural dissociation between computational and subjective image complexity. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000605

Zhu, R., & Meyers-Levy, J. (2005). Distinguishing between the meanings of music: when background music affects product perceptions. *Journal of Marketing Research*, *42*(3), 333-345. http://www.jstor.org/stable/30162377

Zhu, T., Li, L., Chen, P., Wu, J., Yang, Y., & Li, Y. (2024). Emotion-aware hierarchical interaction network for multimodal image aesthetics assessment. *Pattern Recognition*, *154*, Article 110584. https://doi.org/10.1016/j.patcog.2024.110584