Universität
Basel

Fakultät für
Psychologie

The Presentation of Images: Effects on the Viewer's Reaction

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

Claire Marie Sarah Reymond

aus Le Chenit (VD), Schweiz

Basel, 2022

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. Klaus Opwis (Erstgutachter)

Prof. Michael Renner (Zweitgutachter)

Datum des Doktoratsexamen: 4. April 2022

DekanIn der Fakultät für Psychologie

Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

- Reymond, C., Pelowski, M., Opwis, K., Takala, T., & Mekler, E. D. (2020). Aesthetic Evaluation of Digitally Reproduced Art Images. *Frontiers in Psychology*, *11*.
- Reymond, C., Vornhagen, J. B., Pelowski, M., Opwis, K., & Mekler, E. D. (submitted). Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs.
- Reymond, C., Widmer Beierlein, S., Müller, C., Reutimann, R., Kuntner, K. P., Falcon Garcia, N., Grumbinaite, I., Hemm, S., Degen, M., Parillo, F., Karlin, S., Park, S., Renner, M., & Blechschmidt, A. (submitted). Naming Images in Aphasia: Effects of Graphic Representations and Photographs on Naming Performance in People with and without Aphasia.

Basel, 4. April 2022

Claire Reymond

Index

Abstract 6

Introduction 8

Theoretical Background 9

Processing Images 9

Models of Art Processing 9

Factors Influencing Image Processing 11

Measuring Reactions to Art Images 12

Naming Images 14

Image Type 15

Measuring Image Naming 15

Aim of this Thesis 16

Research Questions 16

Summary of Manuscripts 20

Manuscript 1: Aesthetic Evaluation of Digitally Reproduced Art Images 20

Manuscript 2: Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs 23

Manuscript 3: Naming Images in Aphasia: Effects of Graphic Representations and Photographs on Naming Performance in People with and without Aphasia 27

General Discussion 30

Relating the Results of this Thesis to the Practice of Image Making 30

Limitations 32

Conclusion 33

References 34

Acknowledgements 43

Appendix 44

Abstract

Background: Perceiving an image evokes a variety of reactions in the viewer. However, not only what is recognized in an image influences the perceiver, but also how an image is presented impacts the reaction to the image. To contribute to the understanding of the relationship between an image's presentation and how this affects the reaction to it, is the aim of this doctoral thesis. To do so, over the course of three studies presented in the form of three manuscripts, we investigated the effect of a basic image feature (color saturation), the context in which an image is perceived, as well as the effect of image type (whether the image is presented as a photograph or a graphic representation) on the responses to an image.

Methods: In each manuscript, we followed a similar path of investigation: We manipulated a particular aspect of how an image is presented and examined the effects of that specific factor on the viewer's response to the image. In each of the manuscripts presented here, we collected data on different responses to images: from directly measurable dimensions such as naming duration and correct naming, to liking, to more complex aesthetic responses, and to emotions perceived in an image. The design of each study was individualized to properly answer the research question.

Results: The results presented in the three manuscripts show that the color saturation—a basic image feature—and the context in which an image is presented influence responses to an image. More specifically, manipulated color saturation affects the liking as well as specific aesthetic reactions to digitally reproduced paintings, but to a different extent for art experts and lay people. We also showed that the affective interpretation of art photographs seen in the context of other photographs is influenced by the valence of these other photographs. But not only neighboring images lead to a change in judgment; The evaluation dimension also affects the emotion perceived in an image. However, whether an image is presented as photograph or graphic representation does not influence naming performance, especially when both types of images include favorable image properties such as color, texture, and shading.

Conclusion: The results presented in this thesis underline the importance of image presentation, as it influences reactions to images. This is also reflected in the practice of image making, where the focus is not only on the creation of the image itself, but likewise on its presentation. Moreover, our results show that not all the factors investigated and manipulated in our studies influence viewer responses in the way we expected. Similarly, in image practice, the designer often assumes—based on his or her expertise—that the viewer's response can be implicitly predicted. However, the results of our studies show that this is not always the case. Therefore, there is a need to empirically investigate reactions to how an image is presented. Especially in visual

communication, where an unambiguous message of the image is intended, the combination of practical image making and empirical investigations could be of added value.

Keywords: Image processing, Image Features, Pictorial Context, Image Types, Experimental Studies

Introduction

What we see in images emotionally moves us, intellectually challenges us, or sometimes even changes our view of the world. But not only *what* we recognize in an image influences us, but more fundamentally *how* an image is presented has an impact on our response to an image. Color panels applied on a canvas in an artwork by Mark Rothko could have a purely informative and illustrative function, but how and in what context these colors are presented move viewers to tears and evoke complex emotional reactions (Pelowski, 2015).

The aim of this dissertation is to contribute to our understanding of how various of those *how*-factors affect image processing. In accordance with a stimulus-oriented approach (Wassiliwizky & Menninghaus, 2021, p. 439), which assumes that the viewer's aesthetic experience is influenced by how a stimulus appears, this thesis focuses on investigating basic image features, the pictorial context in which an image is seen, and the type of an image. Basic image features describe formal characteristics of an image such as color saturation, symmetry, or complexity. Pictorial context is defined as spatial and temporal environment of an image, which consists of other images. Image type refers to the technique used to create an image and categorizes an image as, for example, a photograph, drawing, or graphic representation. By analyzing these factors, my co-authors and I aim to show the importance of *how* an image appears on the reactions evoked in the perceiver. Complementing this, we also examine how responses to these effects vary as a function of interindividual differences.

The theoretical background of this thesis is structured in two main parts: In the first part, *Processing Images,* the field of empirical aesthetics, a model of art perception as well as a selection of studies examining the encounter with art images are briefly presented. In the second part, *Naming Images*, the theoretical background of images in a clinical context and the use of images in aphasia are briefly described. The first two studies presented in this thesis are based on the theoretical background summarized in the first part and examine reactions to art images. The third study is based on the theoretical background presented in the second part. Concluding the chapter on the theoretical background, the aim of this thesis and the three research questions will be discussed. Subsequently, the results of the three studies are presented separately in the form of three manuscripts. In the last section of this thesis, the findings presented in the manuscripts will be discussed and positioned in a larger context within empirical research on image processing and the practice of image-making.

Processing Images

Empirical research on art perception was firstly conducted and described by Gustav Theodor Fechner in *Vorschule der Ästhetik* (Fechner, 1876) as one of the first experimental disciplines in psychological science. Since these early beginnings and the extension of investigations by Berlyne in the 1970s (Berlyne, 1970, 1972, 1973), the discipline of empirical aesthetics has long been neglected. More recently, however, empirical aesthetics has been established as a field of research in cognitive science, driven in particular by new methods in neuroscience, which have also given rise to the subfield of neuroaesthetics. Building on the humanist tradition, empirical aesthetics attempts to contribute to the revealing of how we interact with artistic stimuli. Investigations in this field are concerned with exploring cognitive, emotional, behavioral, and physiological responses to aesthetic[1] experiences such as in art, literature, theater, or music by using empirical methods from the natural sciences (Wassiliwizky & Menninghaus, 2021)[2].

Models of Art Processing

In recent years, several models of art processing have been proposed in the domain of empirical aesthetics (Chatterjee, 2011; Graf & Landwehr, 2015; Locher et al., 2010; for an overview, see Pelowski et al., 2016). One of the first and most influential is the *model of aesthetic appreciation and aesthetic judgments* by Leder, Belke, Oeberst, and Augustin (2004). The model understands the perception of art as information processing[3] and includes bottom-up as well as top-down phases of art processing. The model proposes the elaboration of five stages that subsequently lead to an aesthetic judgment of the visual art based on the evaluation of the cognitive mastering and an emotional reaction to the image as a by-product of the process.

---

[1] In this thesis, the term *aesthetics* denotes a combination of the richness of sensory perceptions with an evaluative judgment. We thereby rely on the quote of Menninghaus et al. (2019, p. 173): "*aesthetics* (…) entails a special judgmental focus on aspects under consideration that are *subjectively perceived as pleasing to our senses and/or cognitive capacities*".

2 No distinction is made in this thesis between the terms a*rt images* and *images* in general. Both terms are used to refer to all types of intentionally created images.

3 Information processing describes humans as information interpreters, ambiguity solvers, and extrapolators, and thus capable of solving a problem given insufficient information and drawing useful conclusions from given information (Informationsverarbeitung, Dorsch Lexikon der Psychologie).

The *Vienna integrated model of top-down and bottom-up processes in art perception* (VIMAP) proposed by Pelowski et al. (2017) expands on the model of Leder et al. (2004) and provides a differentiated overview of behavioral, cognitive, and emotional responses evoked by the perception of art. The VIMAP (Pelowski et al., 2017) proposes five main processing stages based on the first four stages in the model of Leder et al. (2004) (see Fig. 1 in the Appendix) as well as two later stages. Stage 1 is described as *pre-classification,* a state prior to encountering the artwork. This stage is influenced by several factors that can affect the forthcoming cognitive experience. These are, for example, the context in which the art is seen, the viewer's expectations towards the stimuli, his personality, or his current mood. When the artwork is encountered, stages 2, 3 and 4 are proposed that involve mostly automatic, bottom-up processing. Stage 2, described as *perceptual analysis,* is characterized by a rapid (approximately 100ms) analysis of basic image features. Stage 3 is defined as *implicit memory integration* resulting from the first visual impression of the image. At this stage, basic formal image features are grouped to form visual patterns. The information extracted from the artwork during the previous stages is then further processed on stage 4, the *explicit classification*. The first impression of the image is now associated with the context and the memory of the viewer. At this stage, a deliberate engagement with what is seen takes place: What the image shows can be verbalized, and meaning is assigned to what is seen. Stage 4 is also described as an evaluation of emotion, allowing the viewer to switch to an aesthetic (form focused) or pragmatic (content focused) approach. Stage 5, *cognitive mastery*, is described as a cognitive (top-down) process merging cognitive and affective information. The viewer combines the information gathered in the prior stages and forms a coherent meaning of what he or she sees in the attempt to create evaluations and associations. The phases of the model are described as occurring in a linear succession which may return to earlier phases in the course of processing and whose elements are partially connected by feedback loops.

The most significant addition of the VIMAP (Pelowski et al., 2017) over the model proposed by Leder et al. (2004) is the description of five specific outcomes resulting from the seven processing stages. The different outcomes are based on two checks: A schema congruence check, which involves the expectation of cognitive mastery and a self-relevance check which is described as the viewer's importance to the self. Depending on whether these checks result in a high or low assessment, this will lead to different outcomes. Outcome 1 is characterized as a low arousal state and evokes little emotion. Outcome 2 results in an experience of novelty that can involve pleasure, confusion, or even feelings of sublimity and awe. In outcome 3 the viewer experiences a sense of flow, harmony, and emotional resonance. However, if self-relevance is high and, at the same time, schema congruence is low this will lead to stage 6, *secondary control*. This stage is described as the awareness of the insufficiency of one's schema and an attempt to reduce the discrepancy between high self-relevance and low schema congruence. Processing this stage will result in outcome 4, if the need to cope is low, characterized as negative feelings such as anger, shame, and sadness. But if the need to cope is high, the process will go on to stage 7, *self-knowledge, metacognitive evaluation*, resulting in outcome 5, which denotes a transformative experience accompanied by feelings of epiphany, catharsis, and awe (see Fig. 1 in the Appendix).

Factors Influencing Image Processing

In the encounter with images, a variety of factors influence their evaluation (Pelowski & Specker, 2020). Without claiming to be exhaustive, image-related, context-related, and viewer-related factors form the most important categories of influencing variables. According to Pelowski et al. (2017), factors related to the image itself are processed on the second, third, and fourth stages of the VIMAP. Contextual and viewer-related factors influence image processing throughout the entire process. Studies investigating the effect of these three groups of influencing factors are briefly described in the following sections.

Image-related factors. In the interplay between an image and its viewer, image-related factors play a fundamental role in how art images are experienced. Besides factors concerning the image itself, image-related factors also include basic image features characterized as formal aspects of an image. Image features such as the color palette used in the image (Fedorovskaya et al., 1997), the saturation of color (Camgöz et al., 2002, 2004; Ireland et al., 1992), the contrast observed in an image (Dongen & Zijlmans, 2017), the complexity (Berlyne, 1970; Birkin, 2010; Bradley et al., 2007), or the symmetry of figures (Gartus & Leder, 2013; Leder et al., 2019; McManus, 2005) have been shown to affect the evaluation and reactions to images. These basic image features are processed bottom up and within the first 100ms of encounter with an image (stage 2, *perceptual analysis* of the VIMAP, Pelowski et al., 2017). As such, they can be understood as a primary influencer of the initial and subsequent "continuous affective evaluation" inherent to processing art (Leder et al., 2004, p. 492) and thus directly influencing final evaluations.

Context-related factors. The context in which the viewer sees the image affects image processing on all stages as proposed by the VIMAP (Pelowski et al., 2017). Context can thereby include a variety of aspects (for an overview, see Pelowski & Specker, 2020). The social context in which the viewer finds himself e.g., if the image is perceived alone or together with others (Gerger et al., 2011) affects its evaluation. The physical context, that is, whether the image is seen and evaluated in a museum or gallery, in a laboratory setting, at home, or in an urban context (Gartus et al., 2015; Gartus & Leder, 2014; Pelowski, Forster, et al., 2017) changes the aesthetic response. Likewise, the titling and the knowledge provided on an image (Szubielska et al., 2019; Szubielska & Sztorc, 2019) as well as the presumed knowledge of whether an image is art or not influences its evaluation (Gerger et al., 2014; Pelowski, Gerger, et al., 2017; Wagner et al., 2014).

Viewer-related factors. In the encounter with an image, the viewer is the sensing, acting, and processing agent. To this extent, a multitude of viewer-related factors influence the encounter with an image. The expectations the viewer has to his or her encounter with an image (Gerger et al., 2014; Pelowski, Gerger, et al., 2017), the viewer's current mood (Leder et al., 2004), or his or her personality (Fayn et al., 2015; Silvia et al., 2015; Silvia & Nusbaum, 2011). These factors influence the subsequent processing of the image even before it is seen and on all subsequent stages of the process (VIMAP; Pelowski et al., 2017). A further influence is how important it is to the viewer to interact with an image, to understand and appreciate it (Packer & Ballantyne, 2002). Related to this, the expertise of the viewer is crucial in the process of art image processing. Whereas laypersons have been shown to focus on basic aspects of an image like the colors used and the figures depicted, experts tend to focus on conceptual factors and historical meaning (e.g., Belke et al., 2010; Paasschen et al., 2015; P. J. Silvia, 2013; Weichselbaum et al., 2018).

Measuring Reactions to Art Images

As with all empirical research on latent constructs, a central question for studies in the field of empirical aesthetics lies in the form of response measurement. Reactions to art images have been examined along several dimensions in the past, a few of which are described in the following sections.

Liking. The aim of many studies in the field of empirical aesthetics is to measure hedonic responses as a reaction to an art image. The model proposed by Leder et al. (2004, p. 492) describes one outcome dimension of processing an artwork as "aesthetic judgment," which typically manifests itself in questions concerning how much an artwork is liked. Subsequently, different formulations have been used to assess aesthetic judgements: Liking (Dolese et al., 2005; Gartus & Leder, 2013), aesthetic pleasantness (Arielli, 2012; Vessel et al., 2018), preference (Mullennix et al., 2018), or beauty (Tousignant & Bodner, 2014, 2018; Wagner et al., 2014) which only allow a vague approximation between the measurements, but nevertheless often show a high correlation (for a definition and differentiation of terms like aesthetic judgement, aesthetic appreciation, or aesthetic evaluation etc. see Wassiliwizki and Menninghaus, 2021).

Although liking and related ratings are used in many studies, these dimensions map a rather general reaction and may not be able to capture the fine-grained and subtle responses that may arise in the encounter with an image (Silvia, 2009; Wassiliwizky & Menninghaus, 2021). In addition, especially modern art has qualities that cannot be reduced to impressions of beauty and harmony, as is often assumed in psychological research on aesthetics. Therefore, the reduction to the measurement of liking in empirical aesthetics needs to be questioned. This does not mean that liking cannot be a valuable indicator of a general aesthetic response to a stimulus, but it should be assessed in combination with other dimensions to allow conclusive statements about the encounter with images (e.g., Brieber et al., 2015).

Aesthetic reactions based on the VIMAP outcomes. A variety of responses beyond liking can be measured as reactions to the encounter with an image. As an example, the VIMAP (Pelowski et al., 2017) proposes five outcomes that describe specific affective responses in addition to physiological, behavioral, and artwork appraisal factors. In this respect, evaluations along these dimensions can provide information about the outcome of viewing an image. Outcome 1 is described as "facile" or "default" and may be assessed as feedings of boredom. Outcome 2 can be measured by asking about feelings of novelty and provoking new insights. Being moved may be assessed in outcome 3 as these reactions are considered as evoking harmony, flow, and emotional resonance. Outcome 4 describes negative feelings as a reaction to an image, leading the perceiver to re-classify an image as non-art or to end the interaction with the image. The highest outcome of the VIMAP (Pelowski et al., 2017), outcome 5, describes a transformational process evoked by the image where feelings of awe, flow, and catharsis may be experienced and measured. Interest has been shown to be a basic aesthetic reaction to art (Silvia, 2013) and may be described as a component of several outcomes of the VIMAP (Pelowski et al., 2017, p. 98).

Emotions perceived. If we assume that one of the primary purposes of art (images) is to touch us emotionally (Berlyne, 1973), then the measurement of affective responses to images is an essential concern (for an overview on aesthetic emotions, see Menninghaus et al., 2019). Also, in Leder's model of aesthetic appreciation and aesthetic judgments (2004), "aesthetic emotions" are referred to as the second dimension of response to art. These are often assessed via considerations based on basic positive or negative valence or arousal.

Physiological reactions in general (e.g., Krauss et al., 2019), or tears (Pelowski, 2015), and chills (Silvia & Nusbaum, 2011; Wassiliwizky et al., 2015) in particular, have been operationalized in the past as forms of emotion in response to art. The VIMAP (Pelowski et al., 2017, p. 99) describes outcome 3 as a reaction of emotional resonance towards an image. The model proposes a differentiation between depicted and felt emotion. It thereby refers to the perception of an emotion depicted in the image that can resonate in the viewer. Such a distinction between perceived and felt emotion is also proposed by Gabrielsson (2001). He argues for distinguishing between the emotional response itself and the emotion perceived. Felt emotion is thereby described as a genuine emotional reaction to the aesthetic stimulus. Emotion perceived is characterized as a perceptual-cognitive process that does not necessarily emotionally affect the person. An evaluation of the stimuli regarding the emotion perceived requires the viewers to evaluate the emotions he or she discerns in the stimuli regardless of what he or she currently feels.

Naming Images

Investigating reactions to images is not only a concern in empirical aesthetics, but also in a clinical context. Accordingly, images may not only be examined regarding their effect on aesthetic reactions, but also regarding their usefulness as experimental stimuli. Naming images is a common paradigm and valid method used to investigate unimpaired (e.g., Bates et al., 2003; Glaser, 1992; Kohnert, 2004) as well as impaired language processing (e.g., Cotelli et al., 2007; Cuetos et al., 2005; Howard et al., 1985; Kohn & Goodglass, 1985; Laine et al., 1997). As such, images are attributed an important role in assessing the viewer's language abilities. To diagnose the severeness of aphasia, naming images is a common part of standardized and psychometrically validated language tests (e.g., Aachener Aphasie Test; Huber et al., 1983; Boston Naming Test; Kaplan et al., 2001; Bielefelder Aphasie Screening; Richter et al., 2006). Aphasia defines a language disorder due to a brain injury (Huber et al., 2006) whereby the following areas of language are affected: linguistic restrictions in oral and written language production (linguistic expression and writing) as well as in oral and written language comprehension (auditory comprehension and reading) (Schneider et al., 2012). As aphasia patients have relatively well-preserved cognitive abilities to process images (Brown & Thiessen, 2018), the use of images is not only relevant for diagnosis but also offers wide-ranging possibilities in aphasia treatment (e.g., to train language by naming words presented as images). To name an object depicted in an image, however, the image must first be correctly recognized (Heuer, 2016). Thus, the quality of the image can enhance or limit the ability of aphasia patients to recognize represented terms and concepts and consequently influences naming performance.

Image Type

Brown and Thiessen (2018) consider the image type—along with other beneficial factors like richness of detail shown in the image, composition, and layout of images—to be influential when it comes to assessing naming responses. In this thesis, the term *image type* refers to the technique used to create an image (e.g., drawing, photograph, painting). Each image type can be further differentiated: For example, drawings can vary in terms of their level of generalization (from very detailed as in scientific illustrations to generalized and reduced as in pictograms) or photographs can be colored or black-and-white.

Traditionally, several types of images that imply different qualities are used in aphasia therapy and diagnostics. They are usually characterized as drawn or photographic stimuli. *Drawings* mostly refer to a heterogeneous group of black-and-white or colored line drawings. In line drawings the depicted figure is outlined with a black line and presented on a white background, in colored line drawings the outlines are black and the surfaces are colored. Stimuli characterized as *photographs* include all styles of colored or black-and-white photographic images, showing objects or persons embedded in a scene or cropped and shown on a plain background.

Measuring Image Naming

Compared to aesthetic reactions to images that are mostly latent, measuring naming performance based on terms shown in an image may seem comparably evident. Measuring naming performance has been established as a reliable and valid measure of language processing (Alario et al., 2004) in various fields of neuro- and psycholinguistics. To correctly do so, stimulus material must be linguistically controlled with respect to parameters such as word frequency, familiarity, and word length. Because if the depicted words are not comparable along these dimensions, it cannot be said with certainty what the measured effects reflect. Extending the method of measuring naming performance to measure naming speed (in milliseconds), also called naming latency or naming duration, has been applied in language processing research (Bachoud-Lévi et al., 1998; Bates et al., 2003, 2003; Székely et al., 2005). Naming latency is defined as the measurement of the time between the presentation of a stimulus (a written word, an image, a spoken word, or a sentence) and the start of a spoken response, where onset is defined acoustically (Rastle & Davis, 2002, p. 307).

The correctness in naming a term shown on an image varies and must be defined depending on the scope of the study. Thus, different degrees of correctness can be specified and studied: “absolutely correct” if the response given fits exactly the intended term, “extended correct” if the reaction is a synonym or a diminutive of the requested term, or “not correct” if the term does not fit the presented stimulus (Widmer Beierlein et al., 2021). This differentiation is particularly relevant for studies in the field of aphasia research compared to how correctness is handled in practice.

Aim of this Thesis

Given the importance of images in our media saturated environment this thesis aims to investigate the presentation of images and its effects on reactions to images. The three manuscripts that will be discussed in the following chapters are based on the theoretical background presented in the previous chapter.
Models of art perception provide frameworks for empirical research investigating aesthetic reactions to images (Belke & Leder, 2006). As such, the VIMAP (Pelowski et al., 2017) directly forms the construct upon which the study presented in Manuscript 1 is built on, whereas the investigated factors of Manuscript 2 can also be situated here, but the study does not directly base its research questions on the model. In the first manuscript of this thesis, image-related as well as viewer-related factors were examined and reactions to art images were measured using linking ratings and compared to specific outcomes proposed by the VIMAP (Pelowski et al., 2017). In Manuscript 2 we investigated context-related as well as viewer-related factors and measured emotion perceived (Gabrielsson, 2001) in art photographs. In Manuscript 3 of this thesis, naming correctness and naming latencies (Rastle & Davis, 2002) were measured while terms to be named were depicted in two different image types.

Research Questions

Considering the impact images have on our behavioral, cognitive, and emotional responses, it is not surprising that some research exists on factors influencing these processes (Pelowski & Specker, 2020). The abundance of related research originating from empirical aesthetics as well as related disciplines provided us with a rich foundation on which to conduct further research on the effects of image-related, context- and viewer-related factors on responses to art images (Manuscript 1 & 2).

However, there are still unexplored aspects and therefore we address the following research questions.

Color Saturation and its Impact on the Aesthetic Reaction to Art Images (Manuscript 1)

Many people see a work of art for the first time or even solely as a digital reproduction on a computer screen rather than as an original. This places high demands on the digital reproduction of images itself, especially those that were not produced digitally (e.g., paintings). But fundamentally, questions arise about how basic image features of digitally reproduced images impact the viewer's perception and how these reactions differ from reactions to the originals.

Besides size and materiality, color saturation is one aspect that changes considerably in digitally reproduced art images compared to the original artwork. Color saturation is considered a basic image feature that influences image processing from the first few seconds of encounter. To that day, research on the effect of basic image features on the reactions to images has focused on aspects like symmetry (Gartus & Leder, 2013; Leder et al., 2019; McManus, 2005), contrast (Dongen & Zijlmans, 2017), the complexity of the image (Berlyne, 1970; Birkin, 2010; Bradley et al., 2007), or the effect of color saturation in the image in general (Camgöz et al., 2002, 2004). So far, there has been no research on how color saturation affects the evaluation process of digitally reproduced art images in experts and laypersons. By manipulating color saturation, the first research question of this thesis examines the influence of basic image features on aesthetic responses to images:
RQ 1: *How does color saturation in digitally reproduced paintings affect liking as well as specific aesthetic reactions in expert and lay viewers?*

Pictorial Context and its Influence on the Emotion Perceived in Art Photographs (Manuscript 2)

When we perceive images, they are never seen in isolation. Whether in a museum where images are hung side by side, in a magazine where images are displayed on a double page spread, or on a screen where multiple browser windows show different images simultaneously, images are seen in the context of other images. At the same time, it has been shown that the context in which images are perceived has a powerful impact on the cognitive, emotional, and behavioral response to the image (Pelowski & Specker, 2020). Therefore, when images are viewed in a pictorial context it may affect the emotional reaction to an image. In this thesis, we refer to *pictorial context* as a spatial and temporal environment of an image, which consists of images shown side by side or in a sequence (Cohn, 2013, 2015).

Images viewed in direct neighborhood have been shown to influence the perception of a particular image, increasing or decreasing its aesthetic evaluation (Arielli, 2012; Mullennix et al., 2018; Tousignant & Bodner, 2014, 2018). However, how pictorial context affects the emotional reaction towards an image has not yet been assessed. Considering the ubiquity of contextual influences on the interaction with images, understanding context as neighboring images was the aim of the second research question:

RQ 2: *How does the pictorial context influence the emotion perceived in art photographs?*

Compared to research on images in the field of empirical aesthetics, image-related aspects receive comparably little attention in aphasia research. Although the importance of this research gap has been pointed out repeatedly (Brown & Thiessen, 2018; Heuer, 2016), there are only a few studies to date that have investigated the effect of image type in people with aphasia. This is the aim of the study presented in Manuscript 3.

Image Type and its Effect on Naming Performance in Persons with and without Aphasia (Manuscript 3)

The diagnosis and severity of aphasia as well as the therapy that usually follows are based in large part on naming images. At the same time there is still uncertainty regarding which image type is the most favorable for naming tasks in the context of aphasia (Brown & Thiessen, 2018). The current heterogeneity of available visual stimuli poses a particular difficulty for language therapists in providing appropriate pictorial material to their patients. While few studies have demonstrated effects of image type on naming performances (Benton et al., 1972; Bisiach, 1966), others have not (Corlew & Nation, 1975). Some studies have detected favorable effects of reduced drawings on the production of one-word responses (Ma et al., 2009). Others, investigating effects of image type on communication behavior, have shown that if colored photographs were compared to line drawings and pictograms, photographic stimuli evoked better communication performance (Griffith et al., 2014; Ho et al., 2005). However, there is evidence that specific image features like color, texture, or shading positively affect naming performance of objects in adults without neurological deficits (e.g., Biederman & Cooper, 1992; Johnson et al., 1996; Naor-Raz et al., 2003; Palmer et al., 1981; Therriault et al., 2009).

To this date, there has been hardly any investigation comparing stimuli drawn on computer (referred to here as *graphic representations*) that contain favorable image features such as color, texture, and shading with colored photographs on their effect on naming performance in persons with and without aphasia. The third research question was formulated to shed more light on the effects of image type on naming performance:

RQ 3: *How do graphic representations and photographs affect naming latencies and naming correctness in persons with and without aphasia?*

The results of these tree investigations are presented in the form of three manuscripts. All manuscripts are summarized in the following chapters and can be found as complete manuscripts in the Appendix.

Manuscript 1: Aesthetic Evaluation of Digitally Reproduced Art Images

Aim of the Study and Contribution

Art images are more and more perceived as digital reproductions on a computer screen instead of as originals in a museum or gallery. Thereby it is obvious that digital reproductions of art differ in various aspects from the original artworks. Besides the general image quality, the materiality, or the size of the artwork, the most evident change is in the representation of color. However, the effects of subjectively varying color saturation on the evaluation of digitally reproduced art have not been investigated in detail yet. Therefore, the planned contribution of this study was to compare the evaluation of digital reproductions of expressionist and impressionist paintings manipulated to have a high color saturation vs. a saturation similar to the original. The impact of viewing time and expertise were also investigated, two other aspects that may impact the perception of art in online contexts. The dimensions studied were linked to the VIMAP (Pelowski, et al., 2017).

Methods

A 2 x 2 x 2 quasi-experimental mixed-subject design was used to systematically examine the effects of color saturation (high vs. original) and viewing time (100ms vs. unrestricted viewing time) as within subject factors as well as viewer expertise (art experts vs. laypersons) as between subject factors. The dependent variables were "liking" as well as six specific aesthetic reactions derived from the VIMAP (Pelowski et al., 2017) outcomes. The experiment consisted of two blocks (blocks 1 and 2) presented in the same order for all participants and followed by an art questionnaire. Sixteen digitally reproduced color photographs from impressionistic and expressionistic artworks as well as the same images with increased saturation were used as stimuli. Participants were randomly assigned to one of the two different sets of 16 images (containing eight originals and eight manipulated versions of the same paintings). In block 1, each of the images were shown for 100ms followed by a slider displayed on the screen, asking participants to indicate how much they liked the image. In block 2 the same 16 images were shown for an unlimited period of time. In addition to the liking slider, six dimensions ("being-moved", "insight", "surprise", "boredom", "interest", "confusion") derived from the VIMAP (Pelowski et al., 2017) outcomes were rated. Seventy-two lay art viewers and 75 art experts participated in the experiment.

Figure 1: Example of the stimuli used in the study. On the left the digitally reproduced painting with saturation matching the original artwork and on the right with 60% increase in saturation. *Marie, fille du peuple* by Amadeo Modigliani (1918), online collection of the Kunstmuseum Basel.

Results

A repeated-measures ANOVA revealed a significant main effect for saturation (p = .047, η^2_p = .027), with highly saturated images liked more than the original ones across all participants. No significant main effects were found for either expertise (p = .510, η^2_p = .003) or viewing time (p = .093, η^2_p = .019). Post-hoc analysis within the expertise groups revealed that the effect was driven by a significant saturation effect in the laypersons group (p < .001, η^2_p = .168) whereas there was no such effect in the expert group (p = .412, η^2_p = .009).

Regarding the specific aesthetic outcomes, significant main effects of saturation were observed for "surprise" (p < .001, η^2_p = .324), and for "interest" (p = .016, η^2_p = .039), indicating that the manipulation in saturation affected how surprising and interesting participants found the images. Similarly, saturation manipulation significantly affected "boredom" (p = .001, η^2_p = .076), and "confusion" (p < .001, η^2_p = .087), suggesting that participants found increased-saturation images less boring but more confusing. Significant main effects were also found for expertise on "surprise" (p = .020, η^2_p = .037), "interest," (p = .046, η^2_p = .027), "being moved," (p = .022, η^2_p = .035), and "confusion" (p = .047, η^2_p = .027). Experts not only reported more surprise and interest, and felt more moved, but also more confused.

Furthermore, we used Pearson correlation to investigate the consistency of liking ratings between the very short (100ms) and the unrestricted viewing time for original and manipulated saturation for each expert or layperson. Experts' liking of original and manipulated images correlated less strongly between viewing times ($r_{original} = 0.694$, $p < .001$; $r_{manipulated} = 0.730$, $p < .001$) than laypersons' liking ($r_{original} = 0.831$, $p < .001$; $r_{manipulated} = 0.835$, $p < .001$).

Discussion and Conclusion

The findings in Manuscript 1 show that although a significant main effect was detected, effects due to the manipulation of color saturation were rather small. However, for laypersons, the increase in color saturation led to more positive assessments of the images. More specifically, ratings on interest, confusion, surprise, and boredom were affected—to different extents for experts and laypersons. Laypersons found images with increased saturation more interesting, whereas saturation manipulation did not influence experts' interest for the image. Laypersons experienced images with increased saturation as more moving but experts felt less moved. Experts rated original and manipulated images as boring to a similar degree, whereas laypersons found the original images more boring. Finally, experts rated images with increased saturation as more confusing, whereas, for laypersons, increased saturation did not influence how confused they felt. Furthermore, we showed that laypeople's first impression of a painting is strongly consistent with their preference for the artwork when there is sufficient time to view and evaluate it. At the same time, art experts are not quite as consistent in their judgments during very short perception times compared to unrestricted viewing time. These results are particularly important, not only to show the differences in effects based on the evaluative dimension, but also to show how manipulations of saturation affect perception differently depending on the viewer's experience. Thus, when it comes to digitally reproducing an artwork, increasing or decreasing the color saturation of the digitally reproduced image must be carefully investigated depending on who will look at the image.

Manuscript 2: Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs

Aim of the Study and Contribution

Manuscript 1 showed the effect of color saturation on the aesthetic evaluation of digitally reproduced art paintings. The aim of the study presented in Manuscript 2 was to investigate the pictorial context an image is seen in. As images are never seen alone, it seems obvious that other images seen juxtaposed to a particular image will influence the emotion seen in this image. To investigate this, two related studies were conducted. In Study 1, the emotion perceived in art photographs ("target" images) was compared to when images were presented on their own versus when seen in juxtaposition with negatively or positively valenced non-art ("context") images. Additionally, the influence of the artwork's perceived ambiguity was analyzed. In Study 2, the effect of the perceiver's expertise and the formal similarity between the images were examined as moderator variables on the rated valence of the target image when seen paired with a context image.

Methods

For Study 1 an extended 2x2 experimental design was used with presentation mode (solo vs. paired) and valence of the juxtaposed image (positive vs. negative) as factors. Additionally, two different presentation-order conditions were designed (ambiguity first vs. valence first). Study 1 consisted of two main blocks: in block 1 target images were presented alone and in block 2 target images were presented juxtaposed with context images. The dependent variables were ambiguity and valence of the image in the solo presentation and perceived valence in the paired presentation. Study 1 included 106 participants. The stimuli in Study 1 and 2 consisted of horizontal format photographs depicting landscapes and scenes (See Fig. 2a & 2b). Color photographs of fine art by contemporary Western artists were defined as target images. Context images were selected from the OASIS Image Set (Open Affective Standardized Image Set; Kurdi et al., 2017) and consisted of (non-art) color photographs. Context images used were pre-rated as either having a negative or a positive valence.

Study 2 utilized a 2 x 2 x 2 quasi-experimental design to understand the effects of formal similarity (formal similarity vs. no formal similarity) and valence of the context image (positive vs. negative) as within subject factors and expertise (design experts vs. laypersons) as between subject factors. Study 2 showed target and context images side by side and consisted of two main blocks. In block 1 the formal similarity between the target and context images was rated (See Fig. 2b). In block 2 the perceived emotion of the target image was evaluated. One hundred and seventy-eight persons participated in Study 2, of which 56 persons were defined as experts.

Figure 2a: Example of stimuli presentation in the paired, formally similar, negative condition. Target image on the left: Thomas Keller, *Ohne Titel*, from the series *Häuser – Where Distance Lives*, 2004–2010, www.thkeller.com; paired with a negatively valenced context image that is formally similar on the right; OASIS Image Set, Flood 3.

Figure 2b: Example of stimuli presentation in the paired, formally not similar, negative condition. Target image on the left, context image on the right; OASIS Image Set, Explosion 2.

Results

In Study 1, we calculated Pearson correlation between the ambiguity of the target image and the changes in valence from the presentation of single images to pairwise presentation. But no significant correlation was found (positive condition, $r = 0.026$; negative condition, $r = -0.014$). However, as shown in Figure 3, the order in which participants rated ambiguity and valence resulted in the artworks being rated differently in the paired condition.

In order to analyze change in valence ratings of the target images when they were presented alone compared to with a context image, 95% Highest Density Intervals (HDI) were computed. In Study 1, an overall contrast effect was observed: Target images were rated more negatively when paired with a positively valenced context image (mode = –0.24, HDI_{low} = –0.47, HDI_{high} = –0.01) and more positively when presented with negatively valenced context images (mode = 0.26; HDI_{low} < 0.01, HDI_{high} = 0.52) than when seen alone. However, when taking the rating order (ambiguity first vs. valence first) into account, the effect changed for the valence first condition: There was a slightly negative effect size (mode = –0.05, HDI_{low} = –0.42, HDI_{high} = 0.33) when the target images were presented with negatively valenced context images (assimilation effect). When the context image was positively valenced, the target images were rated slightly less positive in the paired condition compared to the single-image condition (mode = –0.17, HDI_{low} = –0.5, HDI_{high} = 0.16) (contrast effect). In general, a negative effect size indicates a more positive rating when the image is seen alone.

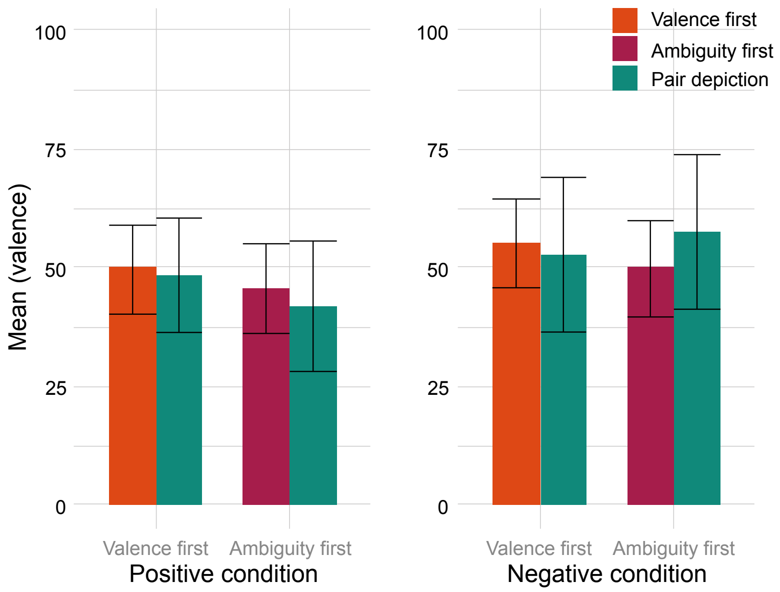


Figure 3: Mean ratings and standard deviation for valence over all stimuli when presented alone in the valence-first condition (orange) and the ambiguity-first condition (red) and when presented paired with a context image (green).

In Study 2, the effect of formal similarity on the perceived emotion of the target images was investigated. Assimilation effects were observed in the negative conditions (target images paired with negatively valenced context images were rated as more negative than when seen alone) and contrast effects were observed in the positive condition (target images paired with positively valenced context images were rated as more negative than when seen alone). Similarly to Study 1, the effect was more pronounced when the target image was seen paired with a negative context image (formally similar condition, mode = –1.57, HDI_{low} = –2.69, HDI_{high} = –0.59; not formally similar condition, mode = –1.84, HDI_{low} = –3.06, HDI_{high} = –0.71) than when the artwork was paired with a positively valenced context image (formally similar condition, mode = –0.56, HDI_{low} = –1.28, HDI_{high} = 0.15; not formally similar condition, mode = –0.82, HDI_{low} = –1.63, HDI_{high} = –0.07). As in Study 1, a negative effect size indicates a more positive rating in the solo condition. Formal similarity between the target and context images did not have a substantial effect.

Discussion and Conclusion

The findings in Manuscript 2 clearly show that neighboring images influence the affective interpretation of art photographs: The emotion perceived in the artwork contrasted away from or assimilated toward the valence perceived in the context image. More specifically, negative images exerted a stronger influence on neighboring images than positive ones did. In Study 1, an overall contrast effect was found: Images shown with a negatively valenced context image were perceived to be more positive than if they were seen alone, whereas images shown with a positive context image were rated more negatively than when rated alone. When the ratings were separated into valence-first and ambiguity-first conditions, a slight tendency for an assimilation effect in the negative condition and a contrast effect in the positive condition was discovered when valence was rated first (Fig. 3). This pattern was also found in Study 2, where the results revealed even larger effects.

The results in Manuscript 2 suggest that if a cognitive context (such as the evaluation dimension) in which an art image is viewed is made salient, it will have an influence on the evaluation of the image. Our results therefore suggest that the evaluative dimension should be considered as part of the pictorial context which contributes to the affective interpretation of an image.

Manuscript 3: Naming Images in Aphasia: Effects of Graphic Representations and Photographs on Naming Performance in People with and without Aphasia

Aim of the Study and Contribution

Manuscript 1 and 2 investigated effects of basic image features and pictorial context on aesthetic reactions in adults without neurological deficits. The goal of this final study presented in Manuscript 3 was to explore the effect of image type in a clinical population and to compare their responses to naming performance in a control group. Naming images is a common paradigm for investigating language processing in persons suffering from aphasia. However, there is uncertainty regarding which image type is the most appropriate for this task. The aim of the third manuscript is therefore to clarify the role of image type on naming performance in persons with and without aphasia. This was done while applying favorable image features like color, texture, and shading to both image types investigated. Naming responses (naming correctness and response latencies) to two image types (graphic representations and photographs) that were created for the study or selected from professional image databases were compared.

Methods

Using a 2 x 2 x 2 quasi-experimental design, naming correctness and naming latencies were measured in eight different conditions: Terms of nouns and verbs depicted as colored photographs vs. as graphic representations as within subject factors and language abilities (persons with aphasia vs. control group of persons without aphasia) as between subject factors. Hundred and twenty-eight images of linguistically controlled German-language terms were developed by professional designers based on photographs. In both image types, the graphic representations as well as the photographs, nouns and verbs were depicted using favorable image features like color, texture, and shading (See Fig. 4). The images were presented in pseudo-randomized sequences on a tablet and all reactions were recorded by a specially programmed application. Thirty-three persons with aphasia (PWA) and 33 age matched participants of the control group (CG) participated in the study.

Figure 4: Depiction of the verb "diving". On the left as photograph (iStockPhoto.com, 487542208, 2015), on the right as graphic representation.

Results

The data was analyzed using generalized linear mixed models (GLMM). Naming differences between persons with aphasia and the control group was significant with participants in the control group naming concepts significantly more correctly than persons with aphasia (Z = -9.210, $p < .001$). There was also a significant difference regarding word class: Nouns were named significantly more correctly than verbs (Z = -5.712, $p < .001$). Our analysis showed no significant difference in naming correctness and latencies between photographic stimuli and graphic representations (Z = -0.012, p = .967) (Fig. 5).

However, when comparing graphic and photographic stimuli in persons with aphasia only, graphic stimuli were named more correctly on average, especially when naming verbs. But this difference was below significance level (See Fig. 5, right column): $M_{Graphic}$ = 56.0, SD = 18.5 vs. M_{Photo} = 53.7, SD = 21.0. The mean of naming verbs correctly was also higher when the terms were presented as graphic representations than if they were depicted as photographs: $M_{Graphic}$ = 49.2, SD = 22.4 vs. M_{Photo} = 48.6, SD = 21.0.

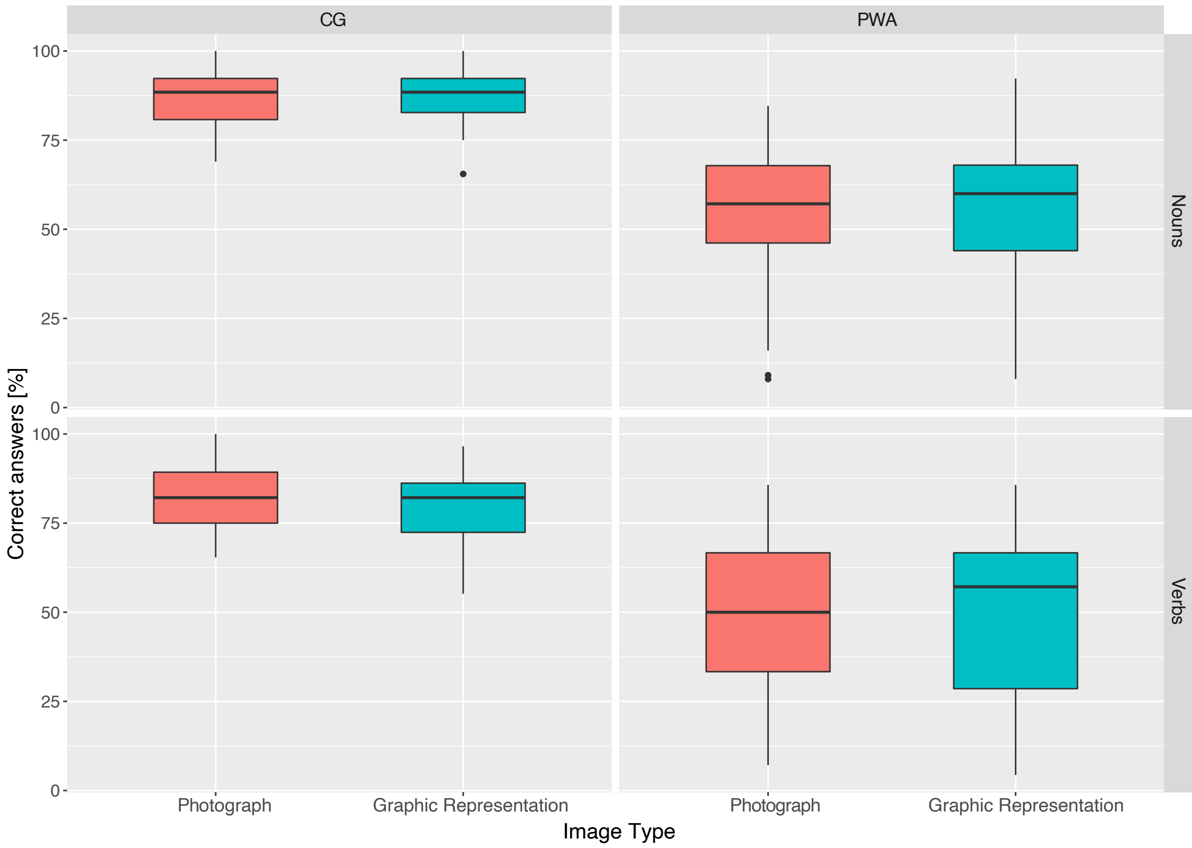


Figure 5: Naming correctness over all eight conditions: photographs vs. graphic representations, nouns vs. verbs, persons with aphasia (PWA) vs. participants in the control group (CG).

Discussion and Conclusion

The two types of images examined in the study presented in Manuscript 3 did not have a significantly different effect on naming performance. Based on the results of Manuscript 3, it was shown that when graphically drawn images are created with features that favor naming performance they can be named as correctly and as fast as photographs. Thus, graphic representations can be used alongside photographs in aphasia therapy and diagnosis. However, our results suggest that naming correctness in persons with aphasia could be supported by depicting verbs as graphic representations. This is especially interesting, as depictions of verbs are underrepresented (for some exeptions see, Akinina et al., 2015; Khwaileh et al., 2018; Spezzano & Radanovic, 2010; Tabak et al., 2010). But further research is needed to evaluate if reduced and translated depictions—defined as graphic representations in our study—could be an advantage for showing verbs to be named in aphasia diagnostic and therapy.

Relating the Results of this Thesis to the Practice of Image Making

Over the course of three manuscripts, three research questions were examined in this thesis. The aim of these investigations was to improve our understanding of how the presentation of an image affects our response to it. We investigated the impact of basic image features, the pictorial context in which an image is seen, and the type of an image on the viewer's reaction.

The first empirical study described in Manuscript 1 provided evidence that a basic image feature such as color saturation has the potential to shape our reaction to an image. As such, the findings substantiate the importance of *how*-factors in the design of images. It is therefore not surprising that the image making process focuses not only on the creation of the image itself, but more importantly on how what is shown is modulated and shaped in an image. Studying the *how* of an image has been a central aspect in training future visual communication designers at the Basel School of Design[4] since the 1960s (e.g., Hofmann, 1985). The emphasis of this tradition lies since then on learning and training practical image making skills and perceptual ability to understand how the meaning of an image can be altered by manipulating formal image features such as color, contrast, or size. More recently, the method of *practice-led iconic research* (Renner, 2010; Renner et al., 2017b, 2017a) has been described as a possibility to systematically investigate how specific image features affect the perceived meaning of an image. This method is based on the approach developed at the Basel School of Design and proposes the creation of image-series in which one image feature at a time is slightly modified. By producing images that differ from each other in only one aspect (e.g., color or size) this approach of practice-based image research allows the designer to practically and subjectively examine—by comparing several images—how an image is changed according to how it is presented.

In Manuscript 2, we examined the effect of the pictorial context on the perceived emotion in an image. Our results clearly showed that images seen in the context influence the affective interpretation of an image, provoking assimilation as well as contrast effects. These results on the effect of pictorial context on the emotion perceived in an image extend findings showing that

[4] Since the 1960s, The Basel School of Design has broadly influenced graphic design and contributed to making Swiss graphic design internationally known as a distinctive style of visual communication.

neighboring images influence the aesthetic evaluation of an image (Arielli, 2012; Mullennix et al., 2018; Tousignant & Bodner, 2018, 2018). These findings demonstrate the importance of consciously designing not only the image itself but also to consider in what context an image is presented. In visual communication as well as in curatorial disciplines (Flacke, 2016; Samara, 2007), time and effort is thus justifiably spent on showing images in an appropriate surrounding because the interpretation of a same image can be modulated by showing it in a new and unseen constellation. This in turn allows the designer or curator to willingly manipulate the context of an image in order to change the intended meaning of an image. Additionally, our results showed that not only do clearly distinguishable factors such as neighboring images influence the viewer's response, but latent constructs such as the evaluation dimension also affect image processing. For the practice of image presentation, this implies that the title of a publication or the name of an exhibition itself can have an influence on the evaluation of an image.

In Manuscript 3, we examined the effect of image type and its impact on naming performance. Our results showed that whether a term was presented as a photograph or as a graphical representation did not affect naming performance significantly. We showed that if an image is made by using basic image features that facilitate recognition and naming (color, shading, texture, etc.), image type has no significant effect on naming performance in individuals with and without neurological deficits. This mirrors results of previous research (Corlew & Nation, 1975) and contrasts with those that show (Benton et al., 1972; Bisiach, 1966) or expect (Heuer, 2016) an effect of naming performance due to different image types. Based on our results, it can be assumed that not the classification into a specific image type has a relevant impact but more precisely the integration of basic image features into the image that favor recognition and naming. Knowing more about the suitability of graphic representations alongside photographic stimuli now allows difficult-to-photograph objects or activities to be created as graphic representations by trained designers.

The Value of Combining Practice based Image Creation and Empirical Investigations on the Perceivers Reaction

In the studies presented here, several aspects of how an image is presented have been shown to influence the viewer's response, while others did not. In Manuscript 2—alongside with examining the pictorial context—we manipulated formal similarity (Arnheim, 1957; Cohn, 2013, 2015) between the images as a factor expected to moderate the juxtaposition effect. But we could not identify any effects due to formal resemblance between the images in our study. On the contrary, formal similarity even showed to be a hindrance in this context. In other words, the pairs of images that were not formally similar influenced each other more than those that were formally similar. This is in contrast to studies showing that similarity as belonging to the same category (Herr et al., 1983) or images from the same painter (Dolese et al., 2005) influenced the effect on neighboring images. Thus, although similarity between images is a premise for contrast and assimilation effects (Fechner, 1876; Mussweiler, 2003), our results suggest that similarity may signify image aspects other than formal relatedness. In Manuscript 3 we investigated the effect of image type on naming performance. Although we expected to find a difference in naming correctness and naming latencies based on different image types (as suggested by studies that have found such differences: Bisiach, 1966; Benton, 1972), we did not find such an effect.

Thus, the manipulation of factors how images are presented can lead to unexpected results, which then open further questions. This highlights the importance of investigating the effects of *how*-factors on the viewer, also in the process of image making. According to Renner (2017a), the practice of image making, and especially practice-led iconic research could be extended through the empirical evaluation of the perceiver's genuine reactions to the images (p. 22). Since it cannot be assumed that reactions to images can be unambiguously anticipated entirely based on practical expertise in image making nor solely on earlier empirical studies, integrating empirical research on viewer reaction into the image-making process could take the creation of images further. In other words, if the perceivers' reaction to the images could be brought back into practical image creation to inform the designer, images that correspond to both the designer's intent and the viewer's understanding could be developed with more precision. This does not mean that all types of images should be created using such an approach. Images that have an unexpected effect on the viewer and thus evoke new, unfamiliar reactions serve an expectation in several artistic contexts. However, a combined investigative approach could be beneficial in visual communication, a field that aims to produce functional and aesthetic visual artifacts. Especially images that are intended to achieve an unambiguous purpose, such as to be used in aphasia diagnostics and therapy, could thus be produced more purposefully.

Limitations

The studies presented here improve our understanding on how specific image aspects influence our reaction towards an image. However, important limitations need to be addressed. A fundamental limitation with which many studies in the field of empirical aesthetics research must deal with is the fact that reactions evoked by images shown on a screen and in an experimental context cannot be fully compared to the experience evoked by an original image in its original context (e.g., Brieber et al., 2015; Gartus & Leder, 2014). In Manuscript 1, we wanted to present a setting that came close to viewing images on a screen, but in Manuscript 2 we aimed to recreate the encounter of images in a pictorial context such as that experienced in a museum. Thus, the results we found in Manuscript 2 should be replicated in front of originals in a museum or gallery context in order to clearly show the importance of pictorial context on the emotional reactions to art photographs.

Similarly, in Manuscript 1, we investigated the effects of saturation manipulation on liking as well as on other specific aesthetic dimensions derived from the VIMAP (Pelowski et al., 2017). However, ratings on liking as well as on those aesthetic dimensions were not pronounced. This could be related to the fact that images seen on a computer screen and in a lab setting generally do not generate as high ratings as images perceived in a museum context (Pelowski et al., 2017). It is therefore possible that an investigation in front of originals could lead to different or more defined results.

Finally, a limitation to our results presented in Manuscript 3 regards the inclusion of favorable image features into our stimuli. As we included favorable image features to representations of nouns and verbs equally, it remains unclear which image features were particularly beneficial to the naming of actions and which were relevant to the naming of objects. Further studies could therefore examine image features (e.g., color, shading, texture) separately to determine their influence on the naming of different word classes.

Conclusion

The results presented in this thesis demonstrate the influence of basic image features, the context in which an image is seen, and the type of image on image processing. Some of these aspects obviously influence the viewer's reactions to an image and other do not. This is especially relevant because images are conceived for a variety of purposes in our image-rich everyday lives. Therefore, it is important to understand the multiple and varied influences of subtle and sometimes even latent image presentation aspects that affect aesthetic and emotional responses or even naming performance, and to consciously integrate this knowledge into the image creation process.

In addition, a connection between empirical research and practice-based image research could add value to future investigations on image processing as well as to the process of image creation. If empirical investigations are provided with meaningful images and with questions relevant to the practice of image making, and empirical research in turn empirically investigates the experience of the viewer, a better understanding of the psychological mechanisms underlying the perception of an image could be gained as well as support the creation of meaningful images.

References

Akinina, Y., Malyutina, S., Ivanova, M., Iskra, E., Mannova, E., & Dragoy, O. (2015). Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*, *47*(3), 691–707. https://doi.org/10.3758/s13428-014-0492-9

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Arielli, E. (2012). Contrast and Assimilation in Aesthetic Judgments of Visual Artworks. *Empirical Studies of the Arts*, *30*(1), 59–74. https://doi.org/10.2190/EM.30.1.f

Arnheim, R. (1957). *Art and visual perception: A psychology of the creative eye*. University of California Press.

Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where Is the Length Effect? A Cross-Linguistic Study of Speech Production. *Journal of Memory and Language*, *39*(3), 331–346. https://doi.org/10.1006/jmla.1998.2572

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Ching Lu, C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. https://doi.org/10.3758/BF03196494

Belke, B., & Leder, H. (2006). Annahmen eines Modells der ästhetischen Erfahrung aus kognitionspsychologischer Perspektive. *Sonderforschungsbereich 626 (Hrsg.): Ästhetische Erfahrung: Gegenstände, Konzepte, Geschichtlichkeit*.

Belke, B., Leder, H., Harsanyi, G., & Carbon, C. C. (2010). When a Picasso is a "Picasso": The entry point in the identification of visual art. *Acta Psychologica*, *133*(2), 191–202. https://doi.org/10.1016/j.actpsy.2009.11.007

Benton, A. L., Smith, K. C., & Lang, M. (1972). Stimulus characteristics and object naming in aphasic patients. *Journal of Communication Disorders*, *5*(1), 19–24. https://doi.org/10.1016/0021-9924(72)90025-1

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*(5), 279–286. https://doi.org/10.3758/BF03212593

Berlyne, D. E. (1972). Ends and means of experimental aesthetics. *Canadian Journal of*

Psychology/Revue Canadienne de Psychologie, *26*(4), 303–325. https://doi.org/10.1037/h0082439

Berlyne, D. E. (1973). Aesthetics and Psychobiology. *Journal of Aesthetics and Art Criticism*, *31*(4), 553–553. https://doi.org/10.2307/429334

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 121.

Birkin, G. (2010). *Aesthetic complexity: Practice and perception in art & design*. Nottingham Trent University (United Kingdom).

Bisiach, E. (1966). Perceptual factors in the pathogenesis of anomia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *2*(1), 90–95. https://doi.org/10.1016/S0010-9452(66)80030-8

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

Brieber, D., Leder, H., & Nadal, M. (2015). The Experience of Art in Museums: An Attempt to Dissociate the Role of Physical Context and Genuineness. *Empirical Studies of the Arts*, *33*(1), 95–105. https://doi.org/10.1177/0276237415570000

Brown, J., & Thiessen, A. (2018). Using images with individuals with aphasia: Current research and clinical trends. *American Journal of Speech-Language Pathology*, *27*, 504–515. https://doi.org/10.1044/2017_AJSLP-16-0190

Camgöz, N., Yener, C., & Güvenç, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Research & Application*, *27*(3), 199–207. https://doi.org/10.1002/col.10051

Camgöz, N., Yener, C., & Güvenç, D. (2004). Effects of hue, saturation, and brightness: Part 2: Attention. *Color Research & Application*, *29*(1), 20–28. https://doi.org/10.1002/col.10214

Chatterjee, A. (2011). Neuroaesthetics: A Coming of Age Story. *Journal of Cognitive Neuroscience*, *23*(1), 53–62. https://doi.org/10.1162/jocn.2010.21457

Cohn, N. (2013). Visual Narrative Structure. *Cognitive Science*, *37*(3), 413–452. https://doi.org/10.1111/cogs.12016

Cohn, N. (2015). Climbing trees and seeing stars: Combinatorial structure in comics and diverse domains. In I. Toivonen, P. Csúri, & E. Van Der Zee (Hrsg.), *Structures in the Mind* (S. 379–392). MIT Press.

Corlew, M. M., & Nation, J. E. (1975). Characteristics of Visual Stimuli and Naming Performance in Aphasic Adults. *Cortex*, *11*(2), 186–191. https://doi.org/10.1016/S0010-9452(75)80043-8

Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arévalo, A., Cappa, S. F., & Padovani, A. (2007). Action and object naming in Parkinson's disease without dementia. *European Journal of Neurology*, *14*(6), 632–637. https://doi.org/10.1111/j.1468-1331.2007.01797.x

Cuetos, F., Gonzalez-Nosti, M., & Martínez, C. (2005). The picture-naming task in the analysis of cognitive deterioration in Alzheimer's disease. *Aphasiology*, *19*(6), 545–557. https://doi.org/10.1080/02687030544000010

Dolese, M. J., Zellner, D. A., Vasserman, M., & Parker, S. (2005). *Categorization affects hedonic contrast in the visual arts* [Data set]. American Psychological Association. https://doi.org/10.1037/e674862010-004

Dongen, N. N. N. van, & Zijlmans, J. (2017). The Science of Art: The Universality of the Law of Contrast. *The American Journal of Psychology*, *130*(3), 283–294. JSTOR. https://doi.org/10.5406/amerjpsyc.130.3.0283

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. https://doi.org/10.3389/fpsyg.2015.01877

Fechner, G. T. (1876). *Vorschule der Aesthetik*. Breitkopf & Härtel.

Fedorovskaya, E. A., de Ridder, H., & Blommaert, F. J. J. (1997). Chroma variations and perceived quality of color images of natural scenes. *Color Research & Application*, *22*(2), 96–110. https://doi.org/10.1002/(SICI)1520-6378(199704)22:2<96::AID-COL5>3.0.CO;2-Z

Flacke, M. (2016). Ausstellen als Narration. In M. Walz (Hrsg.), *Handbuch Museum: Geschichte, Aufgaben, Perspektiven* (S. 253–257). J.B. Metzler. https://doi.org/10.1007/978-3-476-05184-4_56

Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, *5*(1_suppl), 123–147. https://doi.org/10.1177/10298649020050S105

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. https://doi.org/10.1016/j.actpsy.2015.01.005

Gartus, A., & Leder, H. (2013). The Small Step toward Asymmetry: Aesthetic Judgment of Broken Symmetries. *I-Perception*, *4*(5), 361–364. https://doi.org/10.1068/i0588sas

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*(3), 311–320. https://doi.org/10.1037/a0036847

Gerger, G., Leder, H., Faerber, S. J., & Carbon, C.-C. (2011). When the others matter: Context-dependent effects on changes in appreciation of innovativeness. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, *70*(2), 75.

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

Glaser, W. R. (1992). Picture naming. *Cognition*, *42*(1), 61–105. https://doi.org/10.1016/0010-0277(92)90040-O

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. https://doi.org/10.1177/1088868315574978

Griffith, J., Dietz, A., & Weissling, K. (2014). Supporting Narrative Retells for People With Aphasia Using Augmentative and Alternative Communication: Photographs or Line Drawings? Text or No Text? *American Journal of Speech-Language Pathology*, *23*(2), S213–S224. https://doi.org/10.1044/2014_AJSLP-13-0089

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*(4), 323–340. https://doi.org/10.1016/0022-1031(83)90026-4

Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology*, *30*(8), 943–961. https://doi.org/10.1080/02687038.2015.1081138

Ho, K. M., Weiss, S. J., Garrett, K. L., & Lloyd, L. L. (2005). *The Effect of Remnant and Pictographic Books on the Communicative Interaction of Individuals with Gl*. 16.

Hofmann, A. (1985). Thoughts on the Study and Making of Visual Signs, Basle School of Design/Yale School of Art, 1947 to 1985. *Design Quarterly*, *130*, 1–20.

Howard, D., Patterson, K., Franklin, S., Orchard-lisle, V., & Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, *2*(1), 49–80. https://doi.org/10.1080/02643298508252861

Huber, W., Poeck, K., & Springer, L. (2006). *Klinik und Rehabilitation der Aphasie: Eine Einführung* *für Therapeuten, Angehörige und Betroffene*. Thieme.

Huber, W., Poeck, K., Weninger, D., & Willmes, K. (1983). *Aachener Aphasie Test*. Hogrefe.

Ireland, S. R., Warren, Y. M., & Herringer, L. G. (1992). Anxiety and Color Saturation Preference. *Perceptual and Motor Skills*, *75*(2), 545–546. https://doi.org/10.2466/pms.1992.75.2.545

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive Components of Picture Naming. *Psychological Bulletin*, *120*(1), 113–139.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*.

Khwaileh, T., Mustafawi, E., Herbert, R., & Howard, D. (2018). Gulf Arabic nouns and verbs: A standardized set of 319 object pictures and 141 action pictures, with predictors of naming latencies. *Behavior Research Methods*, *50*(6), 2408–2425. https://doi.org/10.3758/s13428-018-1019-6

Kohn, S. E., & Goodglass, H. (1985). Picture-naming in aphasia. *Brain and Language*, *24*(2), 266–283. https://doi.org/10.1016/0093-934X(85)90135-X

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language*, *91*(3), 294–302. https://doi.org/10.1016/j.bandl.2004.04.001

Krauss, L., Ott, C., Opwis, K., Meyer, A., & Gaab, J. (2019). Impact of contextualizing information on aesthetic experience and psychophysiological responses to art in a museum: A naturalistic randomized controlled trial. *Psychology of Aesthetics, Creativity, and the Arts*.

Laine, M., Vuorinen, E., & Rinne, J. O. (1997). Picture naming deficits in vascular dementia and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *19*(1), 126–140. https://doi.org/10.1080/01688639708403842

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology (London, England: 1953)*, *95*(Pt 4), 489–508. https://doi.org/10.1348/0007126042369811

Leder, H., Tinio, P. P. L., Brieber, D., Kröner, T., Jacobsen, T., & Rosenberg, R. (2019). Symmetry Is Not a Universal Law of Beauty. *Empirical Studies of the Arts*, *37*(1), 104–114. https://doi.org/10.1177/0276237418777941

Locher, P., Overbeeke, K., & Wensveen, S. (2010). Aesthetic Interaction: A Framework. *Design Issues*, *26*(2), 70–79. https://doi.org/10.1162/DESI_a_00017

Ma, X., Boyd-Graber, J., Nikolova, S., & Cook, P. R. (2009). Speaking Through Pictures: Images vs. Icons. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and*

Accessibility, 163–170. https://doi.org/10.1145/1639642.1639672

McManus, I. C. (2005). Symmetry and asymmetry in aesthetics and the arts. *European Review*, *13*(S2), 157. https://doi.org/10.1017/S1062798705000736

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

Mullennix, J. W., Kristo, G. M., & Robinet, J. (2018). Effects of Preceding Context on Aesthetic Preference. *Empirical Studies of the Arts*, 027623741880568. https://doi.org/10.1177/0276237418805687

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*(3), 472–489. https://doi.org/10.1037/0033-295X.110.3.472

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is Color an Intrinsic Property of Object Representation? *Perception*, *32*(6), 667–680. https://doi.org/10.1068/p5050

Paasschen, J. van, 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. https://doi.org/10.1371/journal.pone.0134241

Packer, J., & Ballantyne, R. (2002). Motivational Factors and the Visitor Experience: A Comparison of Three Sites. *Curator: The Museum Journal*, *45*(3), 183–198. https://doi.org/10.1111/j.2151-6952.2002.tb00055.x

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In *Attention and performance* (Bd. 9, S. 135–151). Erlbaum.

Pelowski, M. (2015). Tears and transformation: Feeling like crying as an indicator of insightful or “aesthetic” experience with art. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01006

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., Gerger, G., Chetouani, Y., Markey, P. S., & Leder, H. (2017). But Is It really Art? The Classification of Images as “Art”/“Not Art” and Correlation with Appraisal and Viewer Interpersonal Differences. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01729

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*, 160.

Pelowski, M., & Specker, E. (2020). *The General Impact of Context on Aesthetic Experience (The Oxford Handbook of Empirical Aesthetics)*. https://doi.org/10.1093/oxfordhb/9780198824350.013.42

Rastle, K., & Davis, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 307–314. https://doi.org/10.1037/0096-1523.28.2.307

Renner, M. (2010). Practice-Led Iconic Research. *diid, disegno industriale industrial design*, *42–34*, 76–82.

Renner, M., Reymond, C., & Schubbach, A. (2017a). Introduction. *Visible Language*, *51/52*(3.1), 5–7. http://www.visiblelanguagejournal.com/issue/262

Renner, M., Reymond, C., & Schubbach, A. (2017b). Practice-led Iconic Research: Towards a Research Methodology for Visual Communication. *Visible Language. The Journal of Visual Communication Research*, *51*(Issue 3-52.1).

Richter, K., Wittler, M., & Hielscher-Fastabend, M. (2006). *Bielefelder Aphasie Screening (BiAS) – Zur Diagnostik akuter Aphasien*. https://pub.uni-bielefeld.de/record/1857986

Samara, T. (2007). *Design Elements: A Graphic Style Manual*. Rockport Publishers.

Schneider, B., Wehmeyer, M., & Grötzbach, H. (2012). *Aphasie: Wege aus dem Sprachdschungel* (5. Aufl.). Springer.

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

Silvia, P. J. (2013). Interested Experts, Confused Novices: Art Expertise and the Knowledge Emotions. *Empirical Studies of the Arts*, *31*(1), 107–115. https://doi.org/10.2190/EM.31.1.f

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*(4), 376–384. https://doi.org/10.1037/aca0000028

Silvia, P., & Nusbaum, E. (2011). On Personality and Piloerection: Individual Differences in Aesthetic Chills and Other Unusual Aesthetic Experiences. *Psychology of Aesthetics, Creativity, and the Arts*, *5*, 208–214. https://doi.org/10.1037/a0021914

Spezzano, L. C., & Radanovic, M. (2010). Naming abilities: Differentiation between objects and verbs in aphasia. *Dementia & Neuropsychologia*, *4*(4), 287–292. https://doi.org/10.1590/S1980-57642010DN40400006

Székely, A., Damico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arevalo, A., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25. https://doi.org/10.1016/S0010-9452(08)70174-6

Szubielska, M., Szary, A., & Sztorc, A. (2019). The Influence of Listening to Curatorial Information on Emotions Caused by Artworks, and the Assessment of Contemporary Art by Students. Research in the Space of an Art Gallery. *Przegląd Badań Edukacyjnych*, *1*(28), 57–75.

Szubielska, M., & Sztorc, A. (2019). The Influence of Extended Contextual Information Provided in a Contemporary Art Gallery on Aesthetic Experience of Art Faculties Students. *Polish Psychological Bulletin*, 345–351.

Tabak, W., Schreuder, R., & Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon*, *5*(1), 22–46. https://doi.org/10.1075/ml.5.1.02tab

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processing*, *10*(4), 335. https://doi.org/10.1007/s10339-009-0260-4

Tousignant, C., & Bodner, G. E. (2014). Context effects on beauty ratings of photos: Building contrast effects that erode but cannot be knocked down. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 81–86. https://doi.org/10.1037/a0034942

Tousignant, C., & Bodner, G. E. (2018). Context effects on beauty ratings of abstract paintings: Contrast, contrast, everywhere! *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 369–377. https://doi.org/10.1037/aca0000146

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

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

Wassiliwizky, E., & Menninghaus, W. (2021). Why and How Should Cognitive Science Care about Aesthetics? *Trends in Cognitive Sciences*, *25*(6), 437–449. https://doi.org/10.1016/j.tics.2021.03.008

Wassiliwizky, E., Wagner, V., Jacobsen, T., & Menninghaus, W. (2015). Art-elicited chills indicate states of being moved. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 405–416. https://doi.org/10.1037/aca0000023

Weichselbaum, H., Leder, H., & Ansorge, U. (2018). Implicit and Explicit Evaluation of Visual Symmetry as a Function of Art Expertise. *I-Perception*, *9*(2), 204166951876146. https://doi.org/10.1177/2041669518761464

Widmer Beierlein, S., Reymond, C., Kuntner, K. P., Blechschmidt, A., Degen, M., Müller, C., Falcón García, N., Altermatt, S., Elsener, C., Karlin, S., Park, S., Reutimann, R., Parrillo, F., Rickert, E., Bucheli, S., Grumbinaite, I., Jochmann, A., Harvey, M., Lee, J., … Hemm, S. (2021). E-Inclusion – Ein interdisziplinäres, Schweizerisches Aphasie-Forschungsprojekt. *Aphasie und Verwandte Gebiete*.

Acknowledgements

First and foremost, I am deeply grateful to Klaus Opwis, my PhD supervisor, for offering me the opportunity to explore empirical research on image processing and for giving me the space, encouragement, trust, and support to do so. Thank you for your door being always open for discussions and contemplations of paintings whose authorship remains unclear.

I also express my sincere thanks to Michael Renner, my second supervisor, for introducing me to practice-based image research, for countless conversations about images and for many years of fruitful collaboration at the Institute Digital Communication Environments.

My heartfelt thanks go to Elisa D. Mekler for teaching me the art of empirical research and scientific storytelling. But most of all for inspiring and uplifting conversations about images, design, and empirical aesthetics. If it weren't for you, I would not be where I am today. Thank you.

I am also grateful to my co-authors Jan Vornhagen, Matthew Pelowski, Simone Hemm, Sandra Widmer Beierlein, Christine Müller, Ricarda Reutimann, and Ed Vessel for their valuable contribution to the studies described above as well as other collaborations. Thank you for your critical but always encouraging feedback. I also thank Marianne Notter for her support on statistical questions.

I thank my dear colleagues at the Academy of Art and Design FHNW. First and foremost, Paloma Lopez Grüninger, Jenni Schmitt, Susanne Käser and Marion Fink for your friendship, encouragement, and enriching collaboration.

Special thanks go to my doctoral colleagues at the Center for General Psychology and Methodology, Nicolas Scharowski and Beat Vollenwyder, for the joint (thus less lonely) path to the doctorate with all its ups and downs.

I also express my gratitude to Sandra Cortesi, Sarah Newman, Matthew Battles, and Valda Murphy for the privilege of being part of the Harvard community as a visiting researcher.

I am deeply grateful to Claudia Hoffmann for the friendship that binds us and that has carried me not only through my doctoral thesis.

My deepest gratitude goes to my family, Thomas Brunner, Basile and Antonin Reymond with all my love and affection. I am deeply grateful for the life we share together—and for our extraordinary time in Cambridge.

Appendix

Figure 1

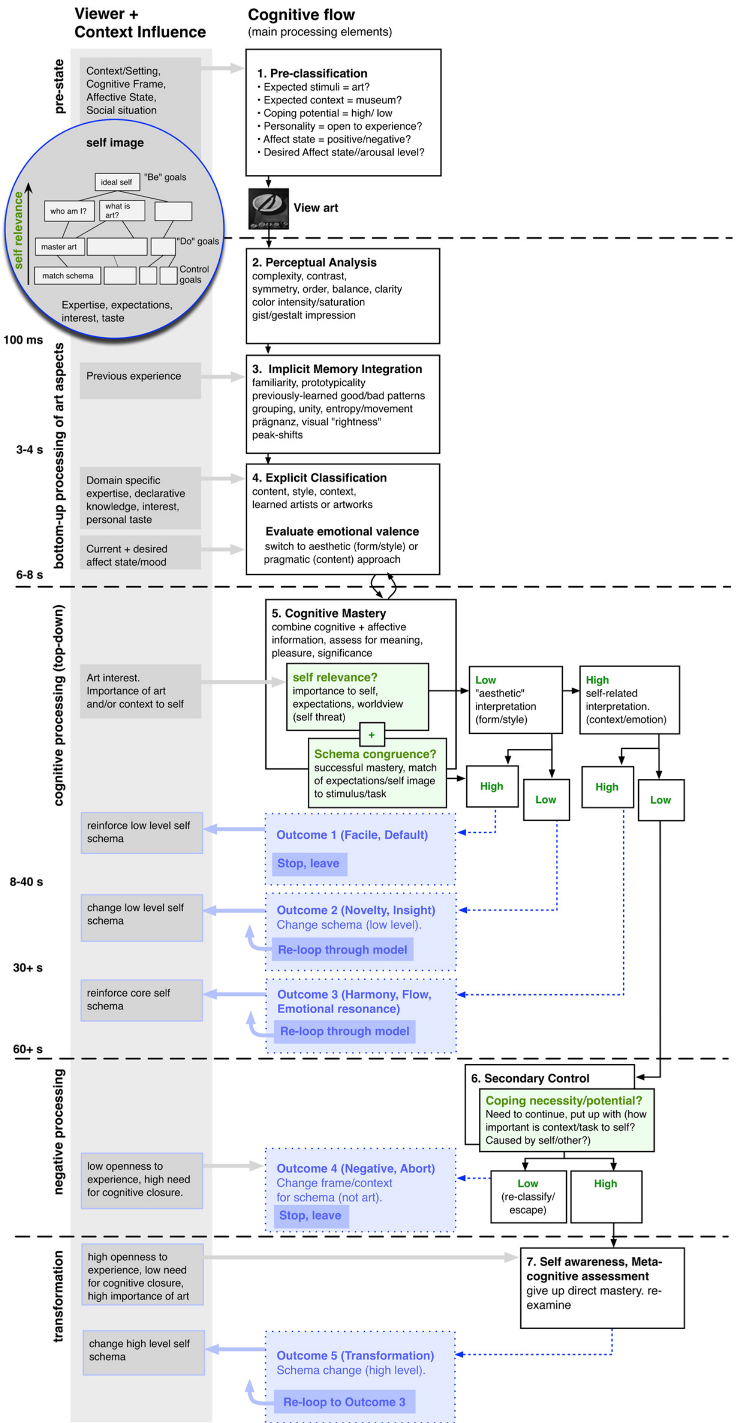


Fig. 1. The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP). Main cognitive model and processing stages is shown in the center (white boxes). Outcomes shown in blue boxes. Leftmost column (in grey) shows personality and self related elements that influence specific stages or processing sequences. Primary opportunities for feedback and re-looping through the model noted by upward facing arrows from output boxes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Complete Manuscripts

Reymond, C., Pelowski, M., Opwis, K., Takala, T., & Mekler, E. D. (2020). Aesthetic Evaluation of Digitally Reproduced Art Images. *Frontiers in Psychology*, *11*.

Reymond, C., Vornhagen, J. B., Pelowski, M., Opwis, K., & Mekler, E. D. (submitted). Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs.

Reymond, C., Widmer Beierlein, S., Müller, C., Reutimann, R., Kuntner, K. P., Falcon Garcia, N., Grumbinaite, I., Hemm, S., Degen, M., Parillo, F., Karlin, S., Park, S., Renner, M., & Blechschmidt, A. (submitted). Naming Images in Aphasia: Effects of Graphic Representations and Photographs on Naming Performance in People with and without Aphasia.

frontiers in Psychology

ORIGINAL RESEARCH
published: 11 December 2020
doi: 10.3389/fpsyg.2020.615575

Aesthetic Evaluation of Digitally Reproduced Art Images

Claire Reymond[1,2]*, Matthew Pelowski[3], Klaus Opwis[1], Tapio Takala[4] and Elisa D. Mekler[4]

[1] Department of Psychology, Center for General Psychology and Methodology, University of Basel, Basel, Switzerland, [2] FHNW Academy of Art and Design, Institute of Visual Communication, University of Applied Sciences and Arts Northwestern Switzerland, Basel, Switzerland, [3] Empirical Visual Aesthetics Lab, Department of Basic Psychological Research and Research Methods, Faculty of Psychology, University of Vienna, Vienna, Austria, [4] Department of Computer Science, Aalto University, Espoo, Finland

Most people encounter art images as digital reproductions on a computer screen instead of as originals in a museum or gallery. With the development of digital technologies, high-resolution artworks can be accessed anywhere and anytime by a large number of viewers. Since these digital images depict the same content and are attributed to the same artist as the original, it is often implicitly assumed that their aesthetic evaluation will be similar. When it comes to the digital reproductions of art, however, it is also obvious that reproductions do differ from the originals in various aspects. Besides image quality, resolution, and format, the most obvious change is in the representation of color. The effects of subjectively varying surface-level image features on art evaluation have not been clearly assessed. To address this gap, we compare the evaluation of digital reproductions of 16 expressionist and impressionist paintings manipulated to have a high color saturation vs. a saturation similar to the original. We also investigate the impact of viewing time (100 ms vs. unrestricted viewing time) and expertise (art experts vs. laypersons), two other aspects that may impact the perception of art in online contexts. Moreover, we link these dimensions to a recent model of aesthetic experience [the Vienna Integrated Model of Top-Down and Bottom-Up Processes in Art Perception (VIMAP)]. Results suggest that color saturation does not exert a major influence on liking. Cognitive and emotional aspects (interest, confusion, surprise, and boredom), however, are affected – to different extents for experts and laypersons. For laypersons, the increase in color saturation led to more positive assessments of an artwork, whereas it resulted in increased confusion for art experts. This insight is particularly important when it comes to reproducing artworks digitally. Depending on the intended use, increasing or decreasing the color saturation of the digitally reproduced image might be most appropriate. We conclude with a discussion of these findings and address the question of why empirical aesthetics requires more precise dimensions to better understand the subtle processes that take place in the perception of today's digitally reproduced art environment.

Keywords: expertise, digitized artworks, subjective aesthetic evaluation, VIMAP, color saturation, evaluation time

Edited by:
Roberto Therón, University of Salamanca, Spain

Reviewed by:
Michal Przemyslaw Muszynski, Université de Genève, Switzerland
Patricia Martin-Rodilla, University of A Coruña, Spain

***Correspondence:**
Claire Reymond
claire.reymond@unibas.ch; claire.reymond@fhnw.ch

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: *09 October 2020*
Accepted: *18 November 2020*
Published: *11 December 2020*

Citation:
Reymond C, Pelowski M, Opwis K, Takala T and Mekler ED (2020) Aesthetic Evaluation of Digitally Reproduced Art Images. Front. Psychol. 11:615575. doi: 10.3389/fpsyg.2020.615575

INTRODUCTION

Perceiving digitally reproduced images has become an indispensable part of our media-suffused everyday life. Digital reproductions specifically play a significant role in people's encounter with works of art. For many viewers, their first encounter with an artwork does not take place in front of the physical artifact in a gallery or museum. Artworks are rather often first perceived as reproductions presented in a digitized form, selected, and compiled by a search engine on the Internet. Such reproductions resemble the original – more or less – and are often thought to be identical to the original artwork. But, digital reproductions of the same painting may differ in several ways, as shown in **Figure 1**: size, quality, resolution, and color can all vary considerably.

In digital reproductions, color is one of the image features expected to differ the most (Strickland et al., 1987; Eschbach and Kolpatzik, 1995; Yang and Rodriguez, 1995; Hu, 2007), especially in web-based presentations of art (see **Figure 1**). According to recent models of art processing (e.g., Leder et al., 2004; Graf and Landwehr, 2015; Pelowski et al., 2017; for an overview, see Pelowski et al., 2016), the color of an image may play a significant role in how art and its online reproductions are experienced. These surface-level features are among the first aspects viewers perceive upon encountering an artwork, are processed in a largely bottom-up fashion, and shape all subsequent stages of the aesthetic experience. Color saturation, in particular, stands out as one of the key components. As one of the three primary dimensions of the human experience of color, along with hue and brightness (Palmer and Schloss, 2015), saturation is processed during the first few milliseconds upon perceiving an image, and therefore likely impacts both our initial assessment – is it beautiful, do I like it? – and all following cognitive and affective processes. This raises the question of whether viewers of digitally reproduced images notice a subjective difference in color saturation, and to what extent this affects the aesthetic assessment of art images viewed online.

To date, only a few studies have investigated the influence of color saturation on the liking of digitized images of art. To address this gap, we provide empirical evidence of the relation between color saturation, expertise, and viewing time. The present study compares high-quality digital color reproductions of impressionist and expressionist works with matching versions of the same paintings with increased saturation. The effect of color saturation is assessed with regard to both general hedonic liking and various aesthetic outcomes tied to later processing stages (Pelowski et al., 2017). Based on the existing studies on the aesthetic evaluation of art (e.g., Leder et al., 2004; Pelowski et al., 2017) and, more specifically, on digitized reproductions of art (e.g., Locher and Dolese, 2004; Siri et al., 2018), we present the results of a quasi-experimental study of 75 art experts and 72 comparably inexperienced viewers of art.

In our results, we demonstrate that color saturation does not exert a major influence on liking. Cognitive and emotional aspects (interest, confusion, surprise, and boredom), however, are affected. As such, our findings suggest that the evaluation of more specific aesthetic reactions is needed to clearly depict the effect visual variables may have on the evaluation of digitized art images. Moreover, we show that color saturation affects viewers differently depending on their expertise in the field of art. While laypersons are positively influenced by increased color saturation, high saturation has a negative influence on experts' evaluations. Also, and in contrast to previous studies (e.g., Kirk et al., 2011; Commare et al., 2018), we show that art experts are less consistent in their aesthetic judgments compared to laypersons. In particular, if they have time to do so, art experts tend to review their initial judgments given after viewing an artwork for a restricted period of time.

Investigating the effect of color saturation in digitized images on the evaluation of art may enable us to further highlight differences in the assessment processes of art experts and laypersons. Furthermore, our findings may encourage art museums and galleries to carefully adjust the color saturation of digitized art images they use on websites or as merchandising products.

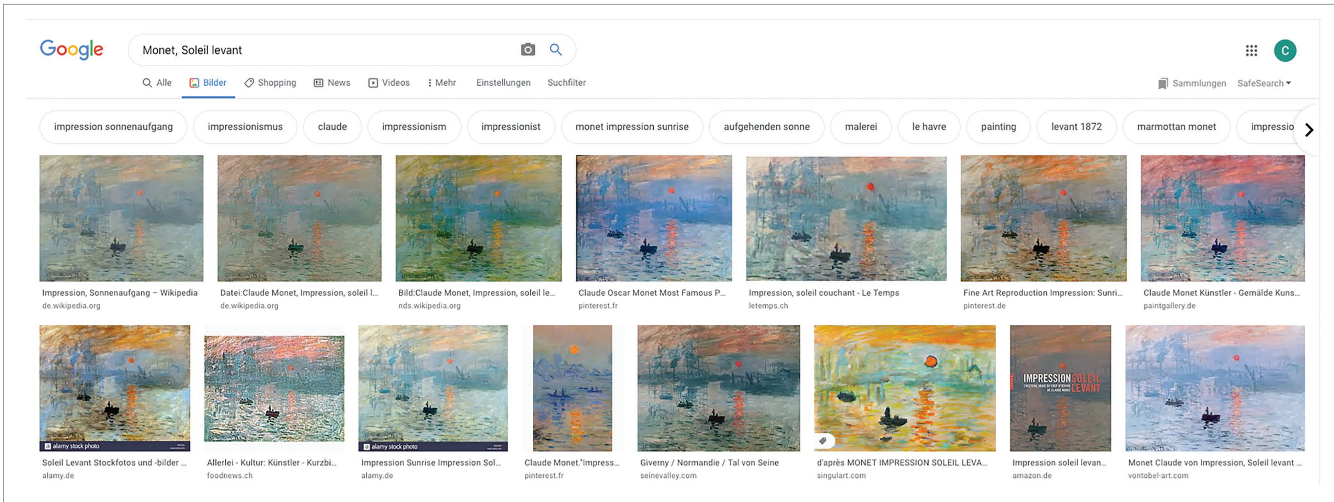


FIGURE 1 | Google search "Monet, Soleil levant" (Accessed July 7, 2020).

RELATED WORK

In the following, we summarize existing research on the aesthetic processing of digitized art images. We also provide an overview of empirical research on color saturation, viewing time, and the relative expertise of the viewer, from which we derived the research questions for our study.

Aesthetic Evaluation of Digitized Art

The differences between original artworks and different reproductions, including digitized art images, have been empirically studied (Berns, 2001; Locher et al., 2001; Locher and Dolese, 2004). Siri et al. (2018), for example, examined both the implicit sensorimotor and explicit cognitive responses of viewers when they observed artworks as originals in their physical form or as high-definition digital reproductions, both within a museum context and presented in the same size. Although there was no visible difference between originals and reproduced images in terms of physiological values, participants explicitly gave higher emotion scores to original artworks than to digital reproductions. In contrast, no significant differences were found with regard to participants' judgment of color intensity and the aesthetic evaluation of digital and original works of art (Siri et al., 2018, p. 217). Locher et al. (1999) compared the perception of three different medial formats of the same artworks: originals, projected slides, and digital images viewed on a computer screen. They found that participants who viewed the reproduced images were aware that they were contemplating a reproduction and focused their attention on the performance and skills of the painter. Moreover, Locher et al. (1999, p. 128) reported that study participants directed their remarks almost exclusively to the art and not to the medium or to the interaction between art and medium. These results specifically show the ability of viewers to adapt to the medium an artwork is presented in Locher et al. (1999, p. 129), therefore, conclude that it is possible to designate "pictorial sameness" between originals and reproduced art images.

Aesthetic Assessment and Color Saturation

When it comes to saturation preference in general, studies have shown that Western adults prefer more saturated colors over less saturated colors, provided the color is not "too vivid" (Granger, 1955, p. 15; see also Valdez and Mehrabian, 1994; Camgöz et al., 2002; Palmer and Schloss, 2015). These preferences lead to an increased attention to colored stimuli (Camgöz et al., 2004; Skulmowski et al., 2016). The perception of color can be described along three primary dimensions (Palmer et al., 2013): hue is characterized as the color's tone, brightness as the lightness of the color, whereas color saturation refers to the relative purity or intensity of the color (Valdez and Mehrabian, 1994; APA Dictionary of Psychology, 2020).

Color Saturation in Digitized Images

Saturated colors are also preferred when it comes to the evaluation of digital images. Tinio and Leder (2009) investigated whether original or manipulated (including low sharpness, low saturation, and low contrast) digital photographs are preferred and showed that higher color saturation is perceived as a characteristic of higher image quality. Fedorovskaya et al. (1997) analyzed the colorfulness of digital images and demonstrated that slightly more colorful images were preferred compared to the original images (Fedorovskaya et al., 1997, p. 110).

The aesthetic quality of digital images affects the attractiveness of websites and subsequently influences viewer behavior (Hong et al., 2004; Li et al., 2014). Regarding the aesthetic evaluation of color on websites, Seckler et al. (2015) showed that blue hues and intermediately to highly saturated colors (together with low complexity and high symmetry) were most preferred. Contrary to this, Skulmowski et al. (2016) found that higher saturation did not lead to a greater preference in website evaluation. They claimed that, depending on the content of the website, color saturation had a negative effect. Skulmowski et al. (2016, p. 386) attribute this effect to the fact that saturated colors are characteristic of rather untrustworthy websites such as those of the yellow press, leading users to perceive very colorful websites as less credible. Reinecke et al. (2013, p. 7) offered a further differentiation by showing that education makes a difference in the assessment of color: "Participants with a doctorate were most negatively affected by high colorfulness, although participants with a higher education preferred websites with a similarly low colorfulness."

Color Saturation and the Aesthetic Evaluation of Art

Models of aesthetic perception (e.g., Leder et al., 2004; Pelowski et al., 2017) describe a rapid initial perceptual analysis during which we perceive, process, and perhaps integrate the surface properties of an image such as color into our general reactions (Seckler et al., 2015; Skulmowski et al., 2016). The color saturation of an artwork can be understood as a primary influencer of the initial and subsequent "continuous affective evaluation" inherent to processing art (Leder et al., 2004, p. 492) and thus often directly influencing final evaluations (Arnheim, 2000; Schloss and Palmer, 2011; Nascimento et al., 2017). Pownall and Graddy (2016) found that the saturation of color significantly increases the price an artwork achieves at an auction. By contrast, studies that have manipulated the lighting of art images (Boust and Ezrati, 2006; Pelowski et al., 2019), which changes the perceived color temperature of an artwork, revealed no difference in aesthetic evaluation.

Due to the mixed evidence, it is unclear to what extent saturation – a feature that can be easily changed in digitally reproduced art images – affects aesthetic evaluation. It is, thus, not yet possible to sufficiently explain the influence of saturation on either the initial assessment or subsequent cognitive and emotional processing of digitized art images.

Aesthetic Evaluation and Time

According to cognitive models of the perception of art (Leder et al., 2004; Pelowski et al., 2017), short-time elaborations take place within the first stage of aesthetic perception and consist of processing surface-level properties of the image as a purely bottom-up visual analysis. A deeper and top-down elaboration of the content and meaning of an image then arguably follows with a longer perception time.

Comparing Short vs. Long Viewing Times

Studies have shown that aesthetic judgments are made very quickly (Verhavert et al., 2018). Research on the aesthetics of websites has revealed that users evaluate the attractiveness of a website within the first 50 ms of encounter and that this rapid evaluation remains consistent over longer perception times (Lindgaard et al., 2006; Tuch et al., 2012; Skulmowski et al., 2016).

According to models of aesthetic perception (Leder et al., 2004; Pelowski et al., 2017), the time an art image is presented should play a role in evaluation, but the effect may only become apparent in later phases of the cognitive and top-down evaluation of the image and would not be discernible during the bottom-up stages of perception (Pelowski et al., 2017). In an examination designed by Smith et al. (2006), participants were presented a work of art for 1 s. Participants remarked afterward that they had barely noticed the image they had seen so briefly. But, bipolar scale-based ratings (e.g., pleasant/unpleasant or simple/complex using the semantic differential technique; Osgood, 1964) from this short viewing time did not significantly differ from ratings with a longer viewing time. Augustin et al. (2008) suspected that even a 50 ms viewing time would be sufficient to become aware of the content of an art image.

Aesthetic Evaluation and Expertise

There is evidence that experts and laypersons differ in their aesthetic judgments (Eysenck, 1972; Winston and Cupchick, 1992; Reinecke et al., 2013; Weichselbaum et al., 2018). "Extensive training (or lack thereof)" (Leder et al., 2019, p. 111) in contemplating, questioning, and creating images exerts an influence on assessing artworks.

Consistency in Aesthetic Evaluation

Recent studies have investigated the consistency of expert judgments. Commare et al. (2018, p. 388) investigated the perceived complexity of artworks in laypersons and art experts. Their results showed that experts were far more consistent in assessing perceived complexity than laypersons when asked to evaluate the complexity of an artwork at two different times. Kirk et al. (2011) investigated the influence of sponsorship on subjective preferences for paintings. They showed that art expertise mitigated the influence of monetary favors in evaluating works of art. In comparison, judgments made by laypersons were favorably influenced by sponsoring, whereas experts' judgments were more consistent with their personal judgments.

Expertise and Viewing Time

Expert and lay judgments are affected by viewing time. Höfel and Jacobsen (2003) showed that laypersons' evaluation of beauty remained consistent over a few days, but this stability decreased with increasing time. According to models of aesthetic perception (Leder et al., 2004; Pelowski et al., 2017), viewers' prior knowledge and expertise impact the cognitive evaluation of an image but hardly play a role in the early stages of perceiving it. By contrast, a recent study by Pelowski et al. (2020) suggests that differences between laypersons and experts already occur at the level of bottom-up processing. They found that persons with greater knowledge of art-like kitsch paintings (which were designed to have bright, highly saturated colors) less when they perceived them for 500 and 6,000 ms compared to when they initially saw them for 100 ms. Experts liked colorful pictures less over longer durations presumably because they switched from focusing on low-level color features to more historically or contextually based assessments. Interestingly, the differences in evaluation between laypersons and experts were already apparent in the 100 ms observation condition.

Aesthetic Reactions Beyond Liking

The target of most studies in the field of empirical aesthetics is to focus on measuring hedonic responses. For example, the model outlined by Leder et al. (2004, p. 492) describes the outcomes of perceiving an artwork as an "aesthetic judgment," which is typically manifested in questions concerning how much an artwork is liked or its relative beauty or quality, all of which often show a high correlation. But there are other possible targets beyond this. Leder also notes "aesthetic emotions," which are often assessed *via* considerations based on basic circumplex models (Russell, 1980) of positive or negative valence or arousal. In assessing the impact that color saturation, viewing time, and expertise might have on judging digital art reproductions, there are also several specific aesthetic reactions that may be empirically observed.

The Vienna Integrated Model of Top-Down and Bottom-Up Processes in Art Perception

The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) proposed by Pelowski et al. (2017) expands on the model of Leder et al. (2004) and further differentiates how the process of aesthetic perception results in distinct experiential outcomes. VIMAP offers a differentiated model of the behavioral, cognitive, and emotional reactions evoked by perceiving art. Five possible outcomes are proposed and can be circumscribed as follows: Outcome 1 is characterized as a low arousal state and evokes little emotion. Outcome 2 consists of an experience of novelty that can result in pleasure, confusion, or even feelings of sublimity and awe. Outcome 3 is when the viewer experiences a sense of flow, harmony, and emotional resonance, whereas in Outcome 4, negative feelings, including anger, shame, and sadness, are more salient. Outcome 5 denotes a transformative experience accompanied by feelings of epiphany, catharsis, and awe. Following these outcomes, it seems that questions such as "Is the image interesting for you?," "Are you touched by the image?," or "Does the image confuse you?" need to be asked in order to differentiate the complex reactions to an image. We, therefore, defined six items for our study that query one characteristic emotional or cognitive dimension of viewing art using the Aesthetic Emotions Scale (Aesthemos; Schindler et al., 2017).

Aesthemos (Schindler et al., 2017) was developed in order to measure manifold reactions to art, film, literature, music, and other art forms with a focus on emotional reactions.

According to VIMAP, the emotional and cognitive elaboration of an artistic stimulus occurs at a later stage than the purely visual evaluation of the stimulus that happens within the first 100 ms of perception. The influence of color saturation on emotional and cognitive reactions to an image should thus be assessed only when the image is processed top-down and can be perceived for an unlimited time. With regard to the relative expertise of the viewer, models of aesthetic perception (Leder et al., 2004; Pelowski et al., 2017) offer little indication of whether expertise or a lack thereof may exert an influence on aesthetic evaluation at particular levels of art perception.

Summary and Research Questions

Digitally reproduced art images are said to be similar to the originals, especially when it comes to specific aesthetic qualities of the image like color intensity, yet a difference is visible in terms of emotional reactions (Locher et al., 1999; Siri et al., 2018). Previous research has shown that higher saturated digitized images are preferred over less saturated ones (Tinio and Leder, 2009). Moreover, color – as a basic cue in visual perception (Palmer and Schloss, 2015) – has a strong impact on whether viewers like art stimuli (Pelowski et al., 2017). As a further factor, the time a visual artifact is perceived influences how it is evaluated (Pelowski et al., 2017). In addition, the level of expertise may modulate the assessment of an artwork (Commare et al., 2018), but this modulation may differ according to the level of processing (Pelowski et al., 2017). Based on previous research (Camgöz et al., 2004; Reinecke et al., 2013; van Dongen and Zijlmans, 2017), we argue that the augmented saturation of digitized reproductions of paintings will increase liking and show an influence on the general variety of outcomes as theorized by VIMAP. This effect should differ between laypersons and art experts. Our research questions can be condensed as follows:

Compared to laypersons, we expect art experts to be less influenced by the manipulation of the image surface (i.e., color saturation), whereas laypersons' liking of digitized images will be influenced by color saturation.

Moreover, the effect of color saturation on liking will be more strongly influenced by viewing time in the case of experts than in the case of laypersons, as predicted by models of the cognitive processing of aesthetic stimuli (Leder et al., 2004; Pelowski et al., 2017).

We further expect that the manipulation on color saturation will be more visible with regard to specific emotional reactions ("boredom," "interest," "insight," "confusion," "being moved," and "surprise") than with regard to liking.

MATERIALS AND METHODS

The study followed a quasi-experimental mixed-subject design. The independent variables were saturation (original vs. manipulated saturation), expertise (art experts vs. laypersons), and viewing time (100 ms vs. unlimited).

Participants

Seventy-two psychology students from the University of Basel [56 female, 15 male, and 1 preferred not to answer; M_{age} = 23.24, *SD* = 5.27, and range = 19–50; all lay viewers of art as assessed *via* post study interviews the Vienna art interest and art knowledge questionnaire (VAIAK); Specker et al., 2020] and 75 art-history students from the University of Basel and the Academy of Art and Design at the University of Applied Sciences and Arts Northwestern Switzerland (47 women, 25 men, and 3 preferred not to answer; M_{age} = 27.25, *SD* = 7.42, and range = 18–62; considered to be relatively expert viewers of art) participated in the experiment. Participants were asked to inform the experimenter about any vision impairments that were not corrigible to normal with glasses or contact lenses. Participants were also asked to inform the experimenter if they had an abnormal color vision. One participant stated that he could only see with one eye and was therefore excluded from the analysis.

Participants were compensated with course credit or monetary compensation (about USD 15). All participants were asked to provide signed informed consent and were informed that they could quit the study at any time and that all data collected in the study would be evaluated anonymously. The study was conducted in accordance with the ethical guidelines of the University of Basel.

Stimuli

The stimuli consisted of 16 high-quality digital color photographs (100 dpi) of paintings from the impressionistic and expressionistic periods of the beginning of the twentieth century, including landscape pictures, portraits, still lives, and groups of figures (see **Supplementary Table A1** for a full list of the artworks, artists, and links to the retrieved paintings). Nine of the original paintings were in the possession of the Kunstmuseum Basel and were downloaded from the online collection of the museum's website.[1] The remaining seven paintings were from other art museums or private collections and were downloaded from the website of the respective museum or auction house. As such, these images were expected to represent the authoritatively most faithful reproductions of the original paintings' contrast and color saturation.[2]

Expressionistic and impressionistic paintings were selected because these styles are known for using color as a formative pictorial element (Alscher, 1968). We deliberately chose original artworks that feature a muted color palette and low color saturation (see **Figure 2** for an example). The fact that high saturation is usually recognized as a distinguishing feature of

[1]https://kunstmuseumbasel.ch/en/collection/collectiononline

[2]The color conformity of the digital reproductions with the original works was checked. The images owned by the Kunstmuseum Basel were evaluated by the authors themselves in terms of color conformity; for the remaining images, the individual museums or auction houses were asked to confirm the color correspondence.

FIGURE 2 | Example of an art image with saturation matching the original and with increased saturation. On the left, the photograph of the painting *Marie* by Modigliani (1918), as it was presented in the online collection of the Kunstmuseum Basel. On the right, the same image with a 60% increase in saturation.

impressionistic and expressionistic style (Alscher, 1968) allowed us to increase color saturation in a way that would not appear overly artificial or incongruous to art experts (Boust et al., 2006).

In addition to the original versions of the 16 images, we also created a matched set of the same 16 paintings but with increased color saturation (leading to a final stimulus set of 32 images). Both the original and manipulated images were scaled to have an identical height of 800 px. **Figure 2** shows a digital photograph of an original artwork and an example of the results of the image manipulation procedure.

All of the manipulations were performed by the first author using Adobe Photoshop CC (version 19.1.4, www.adobe.com) and conducted individually for each image. In consultation with the research team, saturation values were increased between 50 and 80% linearly on the whole image (see **Supplementary Table A2** for more detailed information on the manipulations). This range was chosen to ensure that the changes to color saturation were *clearly visible* at first glance but were also moderate enough to avoid overly colorful-looking images, which risk being perceived as garish (Fedorovskaya et al., 1997). Note that we were specifically interested in viewers' *subjective* responses rather than assessing objective colorimetric thresholds of the saturation level.

Procedure

Participants (psychology students, hereafter "laypersons," and art-history or design students, hereafter "experts") viewed the images in a lab setting, ensuring that they saw the images under the same monitor settings for color and brightness and the same ambient lighting. Five computers were used for the study, arranged in such a way that participants could not see what was on the other screens. The screens of all five computers were uniformly calibrated (using the Apple Display Calibrator Assistant). The brightness of the screens was set to maximum and the automatic adjustment to brightness was switched off. The ceiling lighting always switched on during the experiment. No other light sources were present, except two ceiling windows, which prevented direct light from entering the room. The experiment was programmed in Unipark (2017) and presented on 21.5 in iMac monitors (resolution: 1,920 × 1,080 px, 19.541 × 18.730 in). Viewing distance was about 23.7 in, resulting in a visual angle of ~48°.

Before starting the experiment, participants were requested to give informed consent, to provide demographic information, and to confirm that they had understood the instructions. Moreover, participants were instructed to indicate after completion of the study whether they had recognized any of the images and if so, which ones. However, less than five participants noted that at least one of the images appeared familiar to them. Hence, we did not pursue this further. The study duration was about 30 min on average.

The experiment consisted of two blocks (blocks 1 and 2) presented in the same order for all participants and followed by an art questionnaire. Pilot tests indicated that evaluating the complete set of 2 × 16 images was too taxing, so participants were randomly assigned to one of the two different sets of 16 total images (containing eight originals and manipulated versions of the same paintings). The presentation order of the stimuli was fully randomized. Participants were not informed that image saturation had been manipulated.

Block 1: 100 ms Viewing Time and Liking

Participants were presented with each of the 16 images (in addition to one repeated image shown twice to measure the test-retest reliability; see results). Each image was shown on its own for 100 ms (Pelowski et al., 2017, 2020), centered on a white background. As shown in **Figure 3**, each image was preceded by a fixation cross on a white background for 2,000 ms and followed by a black-and-white noise mask for 1,000 ms. After the noise mask, a slider was displayed on the screen, asking participants to indicate how much they liked the image. Participants were instructed to note their first impression of the image and to indicate their liking as quickly as possible so as to assess their initial response to the image's low-level visual features (i.e., its saturation). This procedure was repeated for all images in the set.

Block 2: Unlimited Viewing Time, Liking, and Specific Aesthetic Reactions

The same 16 images were shown for an unlimited period of time. Again, each image was shown on its own, centered on the top of the screen. Participants were instructed to take their time and to view each painting for as long as they wished. Upon scrolling down, participants had access to the questionnaire. Participants could thus continue viewing the image while answering the questions. Participants again indicated their liking on a rating slider and clicked a "Next" button to access the next image.

Measures

We collected measurements on three different scales during the three phases of the experiment. Blocks 1 and 2 contained

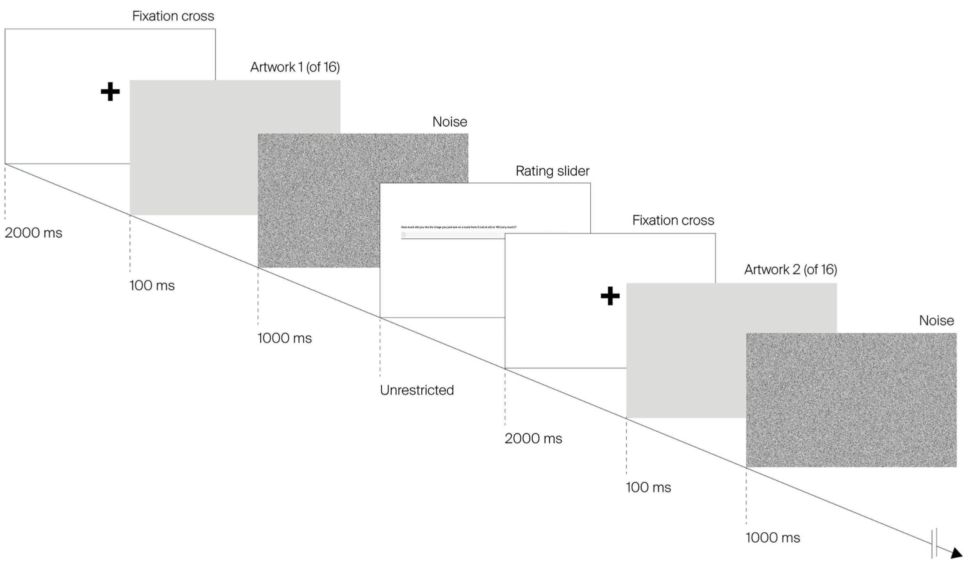


FIGURE 3 | Example sequence of image presentation for block 1 (100 ms viewing time) of the study. Images are cropped for legibility purposes.

a liking slider to rate the stimuli presented. In addition to the liking slider, 12 items to measure six dimensions of the aesthetic evaluation were presented in block 2. During the third phase and before the study was completed, art expertise was assessed.

Liking

To measure participants' liking of an artwork on blocks 1 and 2, a rating slider was displayed below the image offering the possibility to rate the digitized paintings from 0 to 100 (0 = "not at all," 100 = "very much").

Specific Aesthetic Reactions

In block 2, in addition to the liking slider, 12 items [5-point Likert-type scale ranging from "not at all" (1) to "very" (5)] from the Aesthemos (Schindler et al., 2017) were displayed in randomized order under the image. Two items for each of the following six dimensions were defined: "being moved," "boredom," "confusion," "insight," "interest," and "surprise." These dimensions were chosen as they reflect the five outcomes described in the VIMAP (Pelowski et al., 2017). "Boredom" resembles outcome 1 ("facile, default"), "surprise" corresponds to outcome 2 ("novelty, insight"), "being moved" is considered a characteristic of outcome 3 ("harmony, flow, emotional resonance"), "confusion" reflects outcome 4 ("negative, abort"), and "insight" corresponds to outcome 5 ("transformation"). The dimension of "interest" was not attributed to a specific outcome posited by VIMAP, as it is argued to occur on different outcome levels of the model. Yet, it was included to measure a basic reaction to art (Silvia, 2013). Items were reworded into present tense (from past tense in the original Aesthemos questionnaire), as participants were asked to rate their feelings upon viewing the image.

Art-Expertise Questionnaire

Finally, after completing the study, participants were asked to list as many painters as they could name within 60 s, a technique which has been previously employed (e.g., Krauss et al., 2019) to provide a quick estimation of relative art expertise or knowledge. Additionally, participants rated their interest in art *via* the 11-item "interest" battery from the VAIAK (Specker et al., 2020).

RESULTS

All participants completed all portions of the study. The test-retest reliability, as measured by the repeated images in block 1, showed that all image ratings had good reliability (total $r = 0.884$, $p < 0.001$, $n = 147$). Thus, all data were retained for analysis. For all statistical tests throughout the paper, we used an alpha level of 0.05. Due to the exploratory nature of the study, no adjustments for multiple comparisons were performed.

As expected, VAIAK (Specker et al., 2020) scores indicated that experts (art-history and design students) scored significantly higher on art interest ($M = 5.345$, $SD = 0.963$, range = 2.18–6.82) than laypersons [$M = 3.107$, $SD = 0.962$, range = 1–5.36; independent sample t-test $t(145) = 14.094$, $p < 0.001$]. Experts were also able to list significantly more painters within 60 s ($M = 6.45$, $SD = 3.189$, range = 1–14) than laypersons [$M = 4.54$, $SD = 2.222$, range = 1–12; $t(140) = 4.089$, $p < 0.001$].

Impact of Saturation and Expertise on Image Liking for 100 ms and Unlimited Viewing Times

To evaluate the effects of saturation, expertise, and viewing time on image liking, a repeated-measures ANOVA was calculated.

Saturation (original vs. manipulated) × viewing time (100 ms vs. unrestricted) were defined as within-participant factors and expertise (experts vs. laypersons) as a between-participant factor. **Supplementary Table A3** lists descriptive statistics for each image per viewing time and color saturation conditions.

A significant main effect was detected for saturation [$F(1,145) = 3.995$, $p = 0.047$, $\eta^2_p = 0.027$], with more highly saturated images liked more than the original ones across all participants. No significant main effects were found for either expertise [$F(1,145) = 0.436$, $p = 0.510$, $\eta^2_p = 0.003$] or viewing time [$F(1,145) = 2.860$, $p = 0.093$, $\eta^2_p = 0.019$]. The results did reveal, however, a significant interaction between expertise × saturation [$F(1,145) = 10.214$, $p = 0.002$, $\eta^2_p = 0.066$]. Compared to experts (original images: $M = 49.835$, $SD = 14.463$; manipulated images: $M = 49.162$, $SD = 15.034$), laypersons liked more saturated images ($M = 49.382$, $SD = 14.856$) over less saturated images ($M = 46.460$, $SD = 15.172$). **Figure 4** displays mean liking across the eight conditions.

Moreover, we found a significant three-way interaction between saturation × viewing time × expertise [$F(1,145) = 7.636$, $p = 0.006$, $\eta^2_p = 0.050$]. *Post hoc* comparisons within the expertise groups showed that the effect on liking was driven by a significant saturation effect in the laypersons group [$F(1,71) = 14.309$, $p < 0.001$, $\eta^2_p = 0.168$] but not in the experts group [$F(1,74) = 0.682$, $p = 0.412$, $\eta^2_p = 0.009$]. Viewing time had no significant impact on laypersons' liking [$F(1,71) = 1.662$, $p = 0.202$, $\eta^2_p = 0.023$] or on art experts' liking [$F(1,74) = 1.334$, $p = 0.252$, $\eta^2_p = 0.018$].

Consistency in Aesthetic Judgments for Experts and Laypersons Over Different Viewing Times

To analyze the consistency of liking ratings between the very short (100 ms) and the unrestricted viewing time for original and manipulated saturation, we report correlations (Pearson product-moment) between the viewing times for each expert or layperson. Pearson correlation was chosen because image liking was rated on a continuous scale from 0 to 100 (Bortz, 2013). Liking correlations between 100 ms vs. unrestricted viewing times were high overall, indicating that liking ratings remained relatively stable. However, experts' liking of original and manipulated images correlated less strongly between viewing times ($r_{original} = 0.694$, $p < 0.001$, $n = 75$; $r_{manipulated} = 0.730$, $p < 0.001$, $n = 75$) than laypersons' liking ($r_{original} = 0.831$, $p < 0.001$, $n = 72$; $r_{manipulated} = 0.835$, $p < 0.001$, $n = 72$). Overall, both art experts' and laypersons' liking ratings were more consistent for the manipulated images.

To further visualize the relationship between liking ratings, **Figure 5** displays the mean ratings for each participant (individual dots) between the 100 ms and the unrestricted viewing times (y- and x-axis, respectively) for both laypersons and experts (red and blue dots, respectively) and between the original saturation (**Figure 5**, left side) and the manipulated saturation (right side) conditions. Dots above the 45° line indicate that a participant reported higher image liking in the 100 ms condition; dots below the line indicate that participants reported higher liking in the unrestricted condition; those appearing exactly on the line were liked equally across conditions.

Mean ratings over all images show that laypersons were more consistent in their judgments across both viewing times (100 ms: $M = 47.338$, $SD = 15.207$; unrestricted: $M = 48.504$, $SD = 15.089$) than experts, who appeared to increase their liking rating, when there was more time to process the image (100 ms: $M = 48.817$, $SD = 15.863$; unrestricted: $M = 50.180$, $SD = 14.522$). An independent sample t-test revealed a significant difference between the two groups [$t(145) = 0.1331$, $p = 0.042$, $\eta^2_p = 0.011$]. Furthermore, experts took longer to view and evaluate the images in the unrestricted time condition (block 2). Experts spent around 62.44 s (median) per image, whereas laypersons only spent around 44.72 s (median). As requirements for a t-test were not met (variance homogeneity and normal distribution were violated), a nonparametric test was calculated: a Wilcoxon-Mann-Whitney test revealed a significant difference between

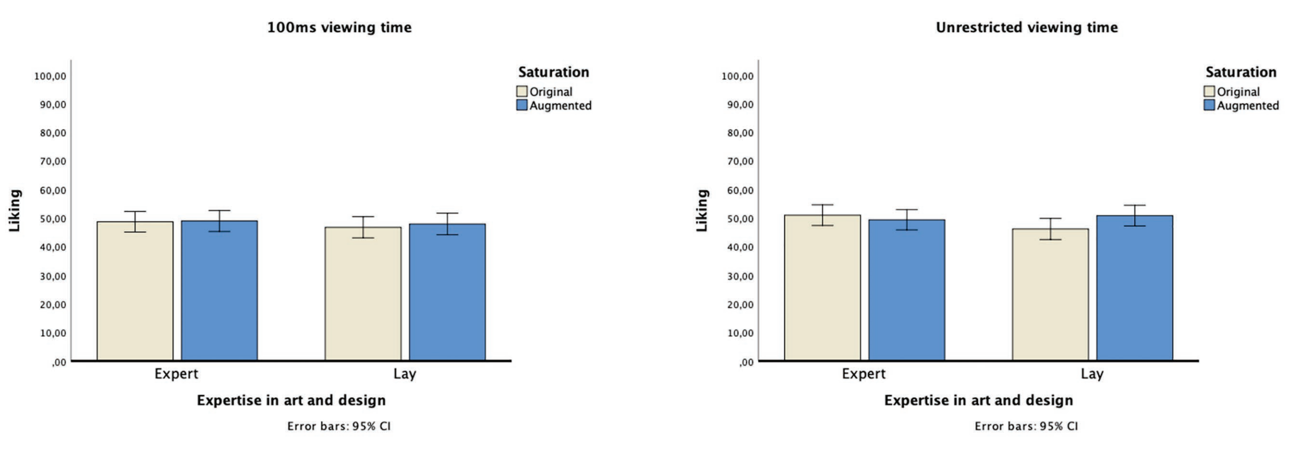


FIGURE 4 | Mean liking for the eight conditions (saturation: original vs. manipulated, viewing time: 100 ms vs. unrestricted, and expertise: experts in art or design vs. laypersons).

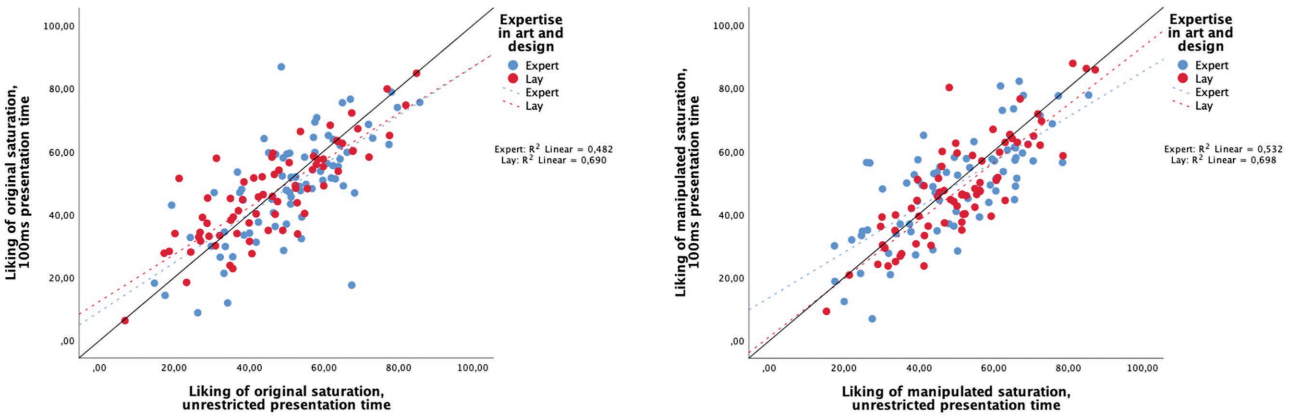


FIGURE 5 | Correlation between image liking and viewing time – 100 ms (*y*-axis) and unrestricted (*x*-axis) – for experts and laypersons. Liking scores for original saturation are on the left and for manipulated saturation on the right.

the groups ($z = -5.33$, $p < 0.001$, $n = 147$), with an effect size of $r = 0.44$, which according to Cohen (1992) corresponds to a large effect.

Saturation, Expertise, and Specific Aesthetic Reactions

To study the effect of saturation in more detail, we investigated six dimensions of aesthetic reactions (block 2) that cover aesthetic assessments beyond basic liking. **Table 1** shows the mean ratings for the six specific aesthetic reactions across all expertise and saturation conditions (see also **Figures 6–8**). To analyze the effects of saturation and expertise, repeated-measures ANOVAs (saturation, within-participant; expertise, between-participant) were calculated for each scale individually.

As listed in **Table 2**, significant main effects of saturation were observed for "surprise" and for "interest," indicating that participants deemed more saturated images more surprising and interesting (see **Figures 6–8**). Similarly, saturation significantly affected "boredom" and "confusion," suggesting that participants found increased-saturation images less boring and confusing. Significant main effects were also found for expertise on "surprise," "interest," "being moved," and "confusion." Experts not only reported more surprise and interest, and felt more moved, but also more confused.

Moreover, significant saturation × expertise interactions were found for all dimensions, except "surprise" and "insight." With regards to "interest," laypersons found images with increased saturation more interesting, whereas experts found them less interesting. A similar pattern was observed for "being moved": laypersons found images with increased saturation more moving, and experts felt less moved. In contrast, with regard to "boredom," experts rated original and manipulated images to a similar degree, whereas laypersons found the original images more boring. Finally, concerning "confusion," experts rated images with increased saturation as more confusing, perhaps because the colors did not match their expectations of the artistic style or of similar art, whereas, for laypersons, increased saturation did not have an effect.

DISCUSSION

An innumerable variety of digitized art images can be found online that may differ substantially from the original in terms of several features, even though they show the same image content and are attributed to the same artist. We conducted a study that reflects a similar encounter with digitally reproduced art. Our focus was to manipulate color saturation (using a matched condition of both high-fidelity versions and increased-saturation versions of the same paintings) – an image component that substantially varies in digital reproductions of art images on the Internet (Eschbach and Kolpatzik, 1995; Yang and Rodriguez, 1995) – and to examine the effect of this surface-level image feature on liking and more specific aesthetic reactions in lay and expert viewers.

The Influence of Saturation on Aesthetic Judgments

We found a main effect for color saturation – more saturated images were liked relatively more when compared within-participants – that extended across both expertise levels and both during short and unrestricted viewing times. The results reflect earlier findings that saturated colors are preferred in general (e.g., Palmer and Schloss, 2015 on saturated colors in general; Seckler et al., 2015 on colors on websites) and underline the argument that the manipulation of saturation exerts an effect on the evaluation of an image. This finding is particularly important when it comes to reproducing art images digitally – which are then to be used, for example, in a virtual gallery or museum, in an online art catalog or as a merchandising material and souvenirs. As our study has shown, an increase in color saturation affects viewers differently depending on their expertise. Since laypersons seem to judge images primarily by their surface texture, increasing the color saturation has a positive effect on their assessment of an artwork. For art experts, who are used to working with images and who focus mainly on the content and meaning of a work, increasing the saturation has the opposite effect and can lead to confusion.

TABLE 1 | Descriptive statistics for specific aesthetic reactions of art experts and laypersons, for original and manipulated saturation.

		Experts		Laypersons		All	
		(*n* = 75)		(*n* = 72)		(*n* = 147)	
Item	Saturation	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
Insight	Original	2.34	(0.59)	2.13	(0.63)	2.24	(0.61)
	Manipulated	2.28	(0.66)	2.14	(0.62)	2.21	(0.64)
Being moved	Original	2.30	(0.66)	1.98	(0.58)	2.15	(0.64)
	Manipulated	2.22	(0.68)	2.08	(0.57)	2.15	(0.63)
Interest	Original	2.85	(0.66)	2.55	(0.61)	2.70	(0.65)
	Manipulated	2.84	(0.67)	2.74	(0.58)	2.79	(0.63)
Confusion	Original	1.80	(0.52)	1.73	(0.54)	1.76	(0.53)
	Manipulated	2.05	(0.68)	1.77	(0.55)	1.91	(0.64)
Surprise	Original	2.14	(0.57)	1.90	(0.56)	2.02	(0.57)
	Manipulated	2.41	(0.60)	2.22	(0.63)	2.31	(0.62)
Boredom	Original	2.19	(0.53)	2.33	(0.69)	2.26	(0.62)
	Manipulated	2.14	(0.55)	2.13	(0.62)	2.14	(0.58)

Item scores range from 1 to 5.

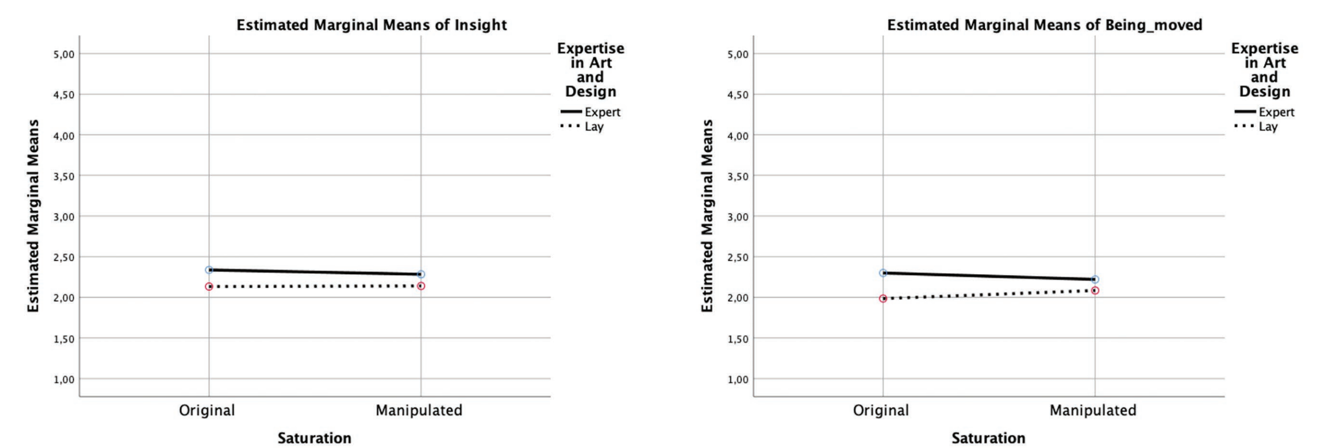


FIGURE 6 | Interaction diagrams for "insight" and "being moved." The diagrams show the direction of the effects for each dimension between the original color saturation and the manipulated color saturation, comparing experts and laypersons.

Our results may also further refine the difference between originals and digital reproductions. Although we used and investigated the aesthetic evaluation of two versions of digitized art images, our results provide more information on the aesthetic evaluation of digital art images. If "faithful high-quality digital reproduction of works of art could be as arousing as the original works of art" (Siri et al., 2018, p. 201), the color saturation of the digitally reproduced artwork must correspond exactly to the original in order not to influence the image's appearance.

At the same time, and against our expectations, our results show that increased saturation had a quite small effect on liking. These results are in contrast to earlier investigations (van Dongen and Zijlmans, 2017) that demonstrated the effect of contrast as a surface-level manipulation on the evaluation of artworks. Perhaps the manipulation of contrast addresses a different level of processing than the manipulation of saturation. While image saturation and contrast are typically subsumed under the same processing level (perceptual analysis; Leder et al., 2004; Pelowski et al., 2017), they are distinct image properties that may have different effects on the liking of an artwork. It should also be noted that participants in our study only liked the images to a moderate extent. In other words, they did not have strong feelings about the images, and saturation manipulation only subtly affected their aesthetic judgments. It remains to be seen to what extent saturation manipulation would impact aesthetic processing for images that viewers strongly like or dislike.

Our findings also differ from studies on the effects of color saturation on the evaluation of websites, which have shown the manipulation of saturation to possess a strong effect (Seckler et al., 2015). It may be argued that art images require a more nuanced and elaborate evaluation than webpages, which might explain the differences in results. Nevertheless, our findings are more in line with the results of Boust and Ezrati (2006), who found that although different lighting conditions alter the color appearance of artworks, viewers' assessment of artworks remained consistent across different color conditions. Boust and Ezrati (2006, p. 6)

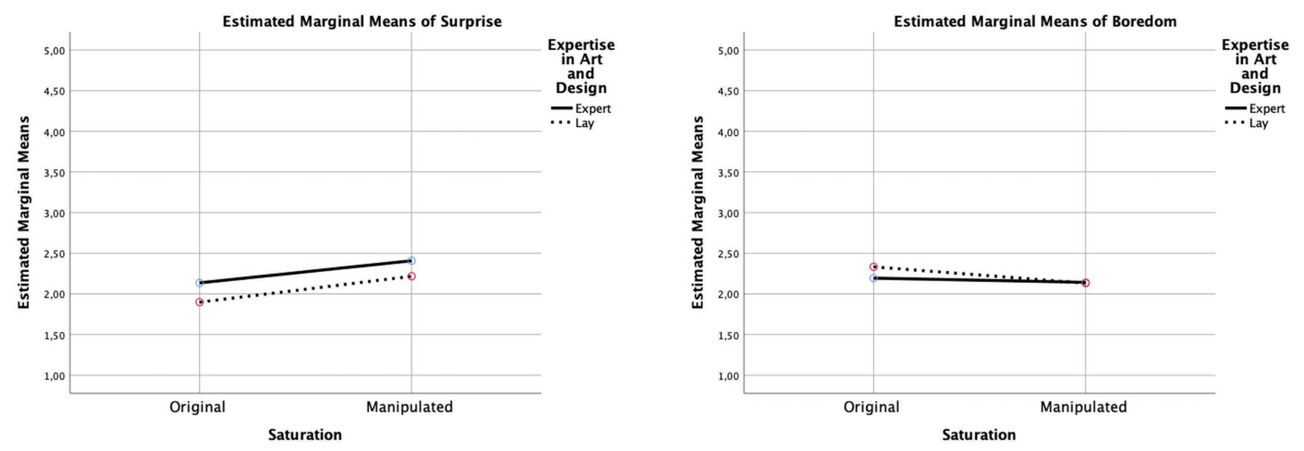


FIGURE 7 | Interaction diagrams for "surprise" and "boredom." The diagrams show the direction of the effects for each dimension between the original color saturation and the manipulated color saturation, comparing experts and laypersons.

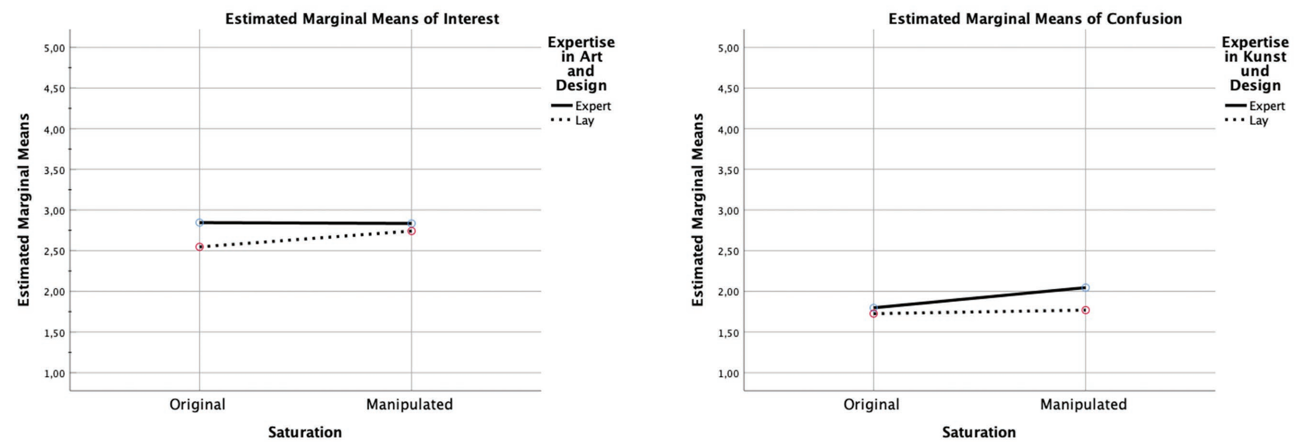


FIGURE 8 | Interaction diagrams for "interest" and "confusion." The diagrams show the direction of the effects for each dimension between the original color saturation and the manipulated color saturation, comparing experts and laypersons.

argue that this may be due to "relational color constancy," suggesting that the relation of colors within the painting is more influential than the absolute value of color. This lack of a substantial effect was also noted by Pelowski et al. (2019), who found very small effects from different color temperatures of lighting on the assessment of artworks.

Effects of Expertise and Time

Differences between our results and previous studies of color saturation in digital images could be related to expertise. For example, Seckler et al. (2015) did not distinguish in their study between web-design experts and laypersons. In our study, although all participants tended to prefer saturated images in general, art experts were comparatively less influenced by manipulations of the image surface, whereas laypersons seemed more susceptible to the colors of an image when indicating how much they liked it. This result itself is in keeping with past studies on the influence of context (e.g., Kirk et al., 2011), which have shown that expertise might tend to insulate against large impacts on appraisals of art from alterations to the image surface. This might also be explained by the relative attention to both low-level surface features of art – including saturation – and more top-down, art-historical aspects, since experts potentially give more emphasis to the latter features when evaluating art (e.g., see Pelowski et al., 2020). Such a result was suggested by the three-way interaction between saturation, expertise, and time in the present study.

Our study also produced interesting findings with regard to viewing time. The lack of difference in liking ratings following both the 100 ms and open-ended viewing duration conditions (and in fact a high correlation between ratings at the level of individual viewers and individual artworks) supports the argument that saturation is one of the features of images that may be processed first, almost immediately following viewing

TABLE 2 | Results of the repeated measures ANOVA on saturation, expertise, and specific aesthetic reactions.

Saturation				Expertise			
Item	***F***	**η^2p**	***p***	**Item**	***F***	**η^2p**	***p***
Insight	0.507	0.003	0.478	Insight	3.177	0.021	0.077
Being moved	0.099	0.001	0.754	Being moved	5.327	0.035	0.022*
Interest	5.918	0.039	0.016*	Interest	4.037	0.027	0.046*
Confusion	13.745	0.087	0.000**	Confusion	4.007	0.027	0.047*
Surprise	69.582	0.324	0.000**	Surprise	5.575	0.037	0.020*
Boredom	11.934	0.076	0.001*	Boredom	0.493	0.003	0.484
Saturation × Expertise							
Item	***F***	**η^2p**	***p***				
Insight	0.856	0.006	0.357				
Being moved	7.405	0.049	0.007*				
Interest	7.265	0.048	0.008*				
Confusion	6.782	0.045	0.010*				
Surprise	0.440	0.003	0.508				
Boredom	4.177	0.028	0.043*				

*Significant at $p < 0.05$; **Significant at $p < 0.001$.

an image. This is in line with the findings of Lindgaard et al. (2006) and suggests that a rapid assessment of visual artifacts not only applies to websites but also to digitized images of paintings when expertise in this field is low. In accordance with cognitive models of aesthetic perception (Leder et al., 2004; Pelowski et al., 2017), a very short viewing time affords only bottom-up perceptual analysis, in which surface-level properties of the image such as saturation are processed and assessments of the image content are not yet included.

Interestingly, one could argue based on our results that, despite its small effect size, saturation may have an impact when considered at the level of basic hedonic (i.e., liking) responses, which may not themselves change or may even inform subsequent analyses. Laypersons' first impression of an image is in strong accordance with their liking of the artwork, even when there is enough time to contemplate and evaluate it. At the same time, as suggested above, our results also show that art experts are not quite as consistent in their liking of judgments in very short vs. unrestricted viewing times. This is in contrast to Commare et al. (2018), who found that experts exhibit more consistency in complexity judgments than laypersons. Our findings support the claim, predicted by the models of Leder et al. (2004) and Pelowski et al. (2017), that expertise affects the evaluation of an image, but only at a later (top-down) stage of processing. Our results suggest that when expertise is low, the assessment of an image at a later stage of processing is consistent with the first impression and the evaluation of the image's surface-level characteristics. But if expertise and background knowledge in this area is more pronounced, it is activated at a later stage and revises the initial, bottom-up visual impression of the perceived image. This is further supported by the fact that experts took significantly more time than laypersons to view and evaluate the images in block 2. This suggests that experts revise their judgments when there is enough time to process the artwork and that they take their time to do so, potentially indicating that experts ground their evaluation more on the content of the image than on surface features, such as saturation. These results are in line with the study conducted by Pelowski et al. (2020), which found that experts might engage in more top-down processes to evaluate an image when there is more time available and thus might reassess low-level features.

Additionally, investigating the perceived visual complexity of the artworks may serve to further differentiate image processing between different levels of expertise and viewing times: for instance, visual complexity and colorfulness have been found to shape viewers' first impressions differently, depending on their age and education level, respectively (Reinecke et al., 2013). Moreover, experts and laypersons perceive the complexity of images differently, with the former tending to appreciate higher perceived complexity more (Reber et al., 2004).

Liking vs. Aesthetic Reactions

The present study supports previous theoretical arguments about aesthetic reactions or features of aesthetic experience beyond basic hedonic liking. As our results show, the effects of saturation manipulation only became apparent when participants were asked to rate their aesthetic reactions to the artworks. Alternatively, our findings reveal that aesthetic reactions indicating an outcome paraphrased by easy-to-achieve positive or also negative sensations on VIMAP showed significant effects. Interestingly, the specific dimensions we looked for in our study showed significant differences according to the manipulation of saturation, and they also revealed differences between experts' and laypersons' evaluations of the images. As shown in **Figures 6–8**, color saturation influenced various aesthetic reactions in laypersons, and they exhibited more differences in their ratings than experts. The dimension "confusion" revealed a crucial aspect: color saturation hardly influenced experts' judgments regarding their interest or boredom while viewing the image and instead led to an irritation when the color intensity was augmented. Experts were more moved by the image, had more interest in it, and experienced a more pronounced sense of insight compared to laypersons. This indicates that, with regard to these aesthetic reactions, art experts cannot be swayed by surface manipulation of an image. Experts did, however, react to a change in color intensity by showing more confusion and surprise.

It should also be noted that the observed effect sizes were rather small. In line with the VIMAP stages of higher-order cognitive processing (Pelowski et al., 2017), we did not anticipate particularly pronounced differences for these dimensions, especially because we expected it to be unlikely that people experience such strong emotional reactions to digital reproductions of art images they see on a computer screen – particularly in a laboratory setting with a sequence of viewed artworks (Pelowski et al., 2017). This explains why both "being moved" and "insight" – which are attributed to outcomes 3 and 5 on the VIMAP, respectively, and are characterized by strong emotional responses and described with feelings of flow or transformation – were not affected by color saturation.

More generally, we argue that in studies of empirical aesthetics, the dimension of liking may not be precise and differentiated enough to adequately reflect and evaluate the perception of art images. In our study, the evaluation of liking could not reflect the diverse dimensions of perceiving an image that were affected by manipulating color saturation, neither in the perception of experts nor in that of laypersons. Including more differentiated and specific factors than liking gave a more detailed impression of the effect of color saturation on the perception of an artwork. We conclude that asking only about the liking of an artwork is not specific enough to evaluate responses to an art image. More research is needed to carefully examine the processes occurring in interaction with art and to analyze specific aesthetic reactions in detail.

Limitations and Future Work

In the following, we address the main limitations of our work and discuss avenues for future work. First, our study employed only a small range of artistic styles (impressionism and expressionism) and a limited pool of viewers. Findings might differ for other image contents or styles, especially if color plays a substantial role in the artwork (e.g., in "Kitsch" artworks, Pelowski et al., 2020). Other aspects of viewers may also have important modulating influences (e.g., Leder et al., 2004). We expect that measuring the current emotional state not only before briefly viewing an image but also before viewing it for an unlimited amount of time could have further contributed to answering the questions in our study. Further limitations to our study may include the homogeneous sample – mostly female participants aged 20–25 took part in our experiment. As it is known that women have a slight preference for pastel colors, a study with more male participants could provide further insights on the effects of color manipulation, given men's preference for saturated colors (Palmer and Schloss, 2015).

Moreover, while comparable to other studies in empirical aesthetics (Locher et al., 1999; Boust and Ezrati, 2006; Siri et al., 2018), the number of images was kept relatively low to minimize participant burden. This trade-off resulted in lower statistical power, which may increase the risk of a Type II error (i.e., false negatives). Future studies should increase their statistical power by using more images and recruiting more participants to assess whether our results can be replicated and whether our study overlooked specific effects.

In line with previous works (Valdez and Mehrabian, 1994; Camgöz et al., 2002; Palmer and Schloss, 2015), we expected that increasing saturation would increase liking. We also assumed that this effect would be more pronounced for laypersons, whereas experts would be more influenced by the content of the image (Commare et al., 2018; Pelowski et al., 2020). That is why we selected images of paintings with a muted color palette and increased their color saturation. For future work, it would be interesting to examine whether images that originally have very saturated colors are liked less when the saturation is reduced.

Next, while we asked participants to indicate at the end of the study whether any of the presented images were familiar to them, we did not measure (perceived) familiarity. Future studies on the effects of color saturation should consider including familiarity as a covariable, as it is a known predictor of image liking (Leder et al., 2004).

In our study, we wanted to investigate the effect of altered saturation as an image feature that often varies unintentionally in digitally reproduced art images on the Internet. In that context, however, images are rarely seen in isolation. The isolated presentation of images in the present study may thus be seen as a limitation in recreating the real situation of how images are seen on the Internet. To investigate the effect that juxtaposed images have on each other is the content of our next study.

In summary, it can be said that whenever a digitized artwork is downloaded from the Internet, the choice of a single version of color saturation out of countless variations exerts an influence on the reception of the image and thus needs to be controlled carefully.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CR, MP, KO, and EM contributed to conceiving and designing the study. TT carried out the measurement of color saturation on the image stimuli. CR prepared the study and collected the survey data. CR and EM performed the statistical analysis. CR, MP, and EM wrote the first draft of the manuscript and wrote the sections of the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Joel Siebenmann for his help in collecting data for this project, Ricarda Reutimann for her support in preparing the data, and Marianne Notter for her feedback on data evaluation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.615575/full#supplementary-material

REFERENCES

Alscher, L. (ed.) (1968). *Lexikon der Kunst, part: 1. A-F*. Leipzig: Seemann.

APA Dictionary of Psychology (2020). Saturation. Available at: https://dictionary.apa.org/saturation (Accessed September 10, 2019).

Arnheim, R. (2000). *Kunst und Sehen. Eine Psychologie des schöpferischen Auges. 3rd Edn*. Berlin: Walter de Gruyter.

Augustin, M. D., Leder, H., Hutzler, F., and Carbon, C. -C. (2008). Style follows content: on the microgenesis of art perception. *Acta Psychol.* 128, 127–138. doi: 10.1016/j.actpsy.2007.11.006

Berns, R. S. (2001). The science of digitizing paintings for color-accurate image archives: a review. *J. Imaging Sci. Technol.* 45, 305–325.

Bortz, J. (2013). *Statistik: Für Sozialwissenschaftler*. Berlin: Springer-Verlag.

Boust, C., Brettel, H., Viénot, F., Berche, S., and Alquié, G. (2006). Color enhancement of digital images by experts and preference judgments by observers. *J. Imaging Sci. Technol.* 50, 1–11. doi: 10.2352/j.imagingsci.technol.(2006)50:1(1)

Boust, C., and Ezrati, J. -J. (2006). Painting viewed under different illuminants: does it change the meaning?

Camgöz, N., Yener, C., and Güvenç, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Res. Appl.* 27, 199–207. doi: 10.1002/col.10051

Camgöz, N., Yener, C., and Güvenç, D. (2004). Effects of hue, saturation, and brightness: part 2: attention. *Color Res. Appl.* 29, 20–28. doi: 10.1002/col.10214

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037//0033-2909.112.1.155

Commare, L., Rosenberg, R., and Leder, H. (2018). More than the sum of its parts: perceiving complexity in painting. *Psychol. Aesthet. Creat. Arts* 12, 380–391. doi: 10.1037/aca0000186

Eschbach, R., and Kolpatzik, B. W. (1995). Image-dependent color saturation correction in a natural scene pictorial image.

Eysenck, H. J. (1972). Personal preferences, aesthetic sensitivity and personality in trained and untrained subjects. *J. Pers.* 40, 544–557. doi: 10.1111/j.1467-6494.1972.tb00079.x

Fedorovskaya, E. A., de Ridder, H., and Blommaert, F. J. J. (1997). Chroma variations and perceived quality of color images of natural scenes. *Color Res. Appl.* 22, 96–110. doi:10.1002/(sici)1520-6378(199704)22:2<96::aid-col5>3.0.co;2-z

Graf, L. K. M., and Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking. *Personal. Soc. Psychol. Rev.* 19, 395–410. doi: 10.1177/1088868315574978

Granger, G. W. (1955). An experimental study of color preferences. ProQuest. Available at: https://search.proquest.com/docview/1290508348?pq-origsite=gscholar (Accessed May 14, 2018).

Höfel, L., and Jacobsen, T. (2003). Temporal stability and consistency of aesthetic judgments of beauty of formal graphic patterns. *Percept. Mot. Skills* 96, 30–32. doi: 10.2466/pms.2003.96.1.30

Hong, W., Thong, J. Y. L., and Tam, K. Y. (2004). Designing product listing pages on e-commerce websites: an examination of presentation mode and information format. *Int. J. Hum Comput Stud.* 61, 481–503. doi: 10.1016/j.ijhcs.2004.01.006

Hu, S. C. -F. (2007). Simple and robust color saturation adjustment for digital images. United States Patent No. US7262780B2. Available at: https://patents.google.com/patent/US7262780B2/en (Accessed March 12, 2020).

Kirk, U., Harvey, A., and Montague, P. R. (2011). Domain expertise insulates against judgment bias by monetary favors through a modulation of ventromedial prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10332–10336. doi: 10.1073/pnas.1019332108

Krauss, L., Ott, C., Opwis, K., Meyer, A., and Gaab, J. (2019). Impact of contextualizing information on aesthetic experience and psychophysiological responses to art in a museum: a naturalistic randomized controlled trial. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000280 (in press).

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., Tinio, P. P. L., Brieber, D., Kröner, T., Jacobsen, T., and Rosenberg, R. (2019). Symmetry is not a universal law of beauty. *Empir. Stud. Arts* 37, 104–114. doi: 10.1177/0276237418777941

Li, X., Wang, M., and Chen, Y. (2014). "The impact of product photo on online consumer purchase intention: an image-processing enabled empirical study" in *Proceedings of the 19th Pacific Asia Conference on Information Systems (PACIS 2014)*; June 24-28, 2014; Association for Information Systems (AIS).

Lindgaard, G., Fernandes, G., Dudek, C., and Brown, J. (2006). Attention web designers: you have 50 milliseconds to make a good first impression! *Behav. Inform. Technol.* 25, 115–126. doi: 10.1080/01449290500330448

Locher, P., and Dolese, M. (2004). A comparison of the perceived pictorial and aesthetic qualities of original paintings and their postcard images. *Empir. Stud. Arts* 22, 129–142. doi: 10.2190/EQTC-09LF-JRHA-XKJT

Locher, P., Smith, L., and Smith, J. (1999). Original paintings versus slide and computer reproductions: a comparison of viewer responses. *Empir. Stud. Arts* 17, 121–129. doi: 10.2190/R1WN-TAF2-376D-EFUH

Locher, P. J., Smith, J. K., and 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, 449–465. doi: 10.1068/p3008

Modigliani, A. (1918). Marie, fille du peuple. Kunstmuseum Basel, Sammlung Online. Available at: http://sammlungonline.kunstmuseumbasel.ch/eMuseumPlus (Accessed September 2018).

Nascimento, S. M. C., Linhares, J. M. M., Montagner, C., João, C. A. R., Amano, K., Alfaro, C., et al. (2017). The colors of paintings and viewers' preferences. *Vis. Res.* 130, 76–84. doi: 10.1016/j.visres.2016.11.006

Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures1. *Am. Anthropol.* 66, 171–200. doi: 10.1525/aa.1964.66.3.02a00880

Palmer, S. E., and Schloss, K. B. (2015). "Color preference" in *Encyclopedia of color science and technology*. ed. R. Luo (Berlin: Springer), 1–7.

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

Pelowski, M., Cabbai, G., Brinkmann, H., Mikuni, J., Hegelmaier, L., and Leder, H. (2020). The kitsch switch—or (when) do experts hate Kinkade art? A study of time-based evaluation changes in top-down versus bottom-up assessment. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000302 (in press).

Pelowski, M., Graser, A., Specker, E., Forster, M., von Hinüber, J., and Leder, H. (2019). Does gallery lighting really have an impact on appreciation of art? An ecologically valid study of lighting changes and the assessment and emotional experience with representational and abstract paintings. *Front. Psychol.* 10:2148. doi: 10.3389/fpsyg.2019.02148

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., and 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. *Phys Life Rev* 21, 80–125. doi: 10.1016/j.plrev.2017.02.003

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

Pownall, R. A. J., and Graddy, K. (2016). Pricing color intensity and lightness in contemporary art auctions. *Res. Econ.* 70, 412–420. doi: 10.1016/j.rie.2016.06.007

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personal. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3

Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., et al. (2013). "Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness." in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Paris, France; April 27–May 2, 2013; 2049–2058.

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161.

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., et al. (2017). Measuring aesthetic emotions: a review of the literature and a new assessment tool. *PLoS One* 12:e0178899. doi: 10.1371/journal.pone.0178899

Schloss, K. B., and Palmer, S. E. (2011). Aesthetic response to color combinations: preference, harmony, and similarity. *Atten. Percept. Psychophys.* 73, 551–571. doi: 10.3758/s13414-010-0027-0

Seckler, M., Opwis, K., and 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. *Comput. Hum. Behav.* 49, 375–389. doi: 10.1016/j.chb.2015.02.056

Silvia, P. J. (2013). Interested experts, confused novices: art expertise and the knowledge emotions. *Empir. Stud. Arts* 31, 107–115. doi: 10.2190/EM.31.1.f

Siri, F., Ferroni, F., Ardizzi, M., Kolesnikova, A., Beccaria, M., Rocci, B., et al. (2018). "Chapter 10: Behavioral and autonomic responses to real and digital reproductions of works of art" in *Progress in brain research*. *Vol. 237*. eds. J. F. Christensen and A. Gomila (Amsterdam: Elsevier), 201–221.

Skulmowski, A., Augustin, Y., Pradel, S., Nebel, S., Schneider, S., and Rey, G. D. (2016). The negative impact of saturation on website trustworthiness and appeal: a temporal model of aesthetic website perception. *Comput. Hum. Behav.* 61, 386–393. doi: 10.1016/j.chb.2016.03.054

Smith, L. F., Bousquet, S. G., Chang, G., and Smith, J. K. (2006). Effects of time and information on perception of art. *Empir. Stud. Arts* 24, 229–242. doi: 10.2190/DJM0-QBDW-03V7-BLRM

Specker, E., Forster, M., Brinkmann, H., Boddy, J., Pelowski, M., Rosenberg, R., et al. (2020). The vienna art interest and art knowledge questionnaire (VAIAK): a unified and validated measure of art interest and art knowledge. *Psychol. Aesthet. Creat. Arts*. 14, 172–185. doi: 10.1037/aca0000205

Strickland, R. N., Kim, C.-S., and McDonnell, W. F. (1987). Digital color image enhancement based on the saturation component. *Opt. Eng.* 26:267609. doi: 10.1117/12.7974125

Tinio, P., and Leder, H. (2009). Natural scenes are indeed preferred, but image quality might have the last word. *Psychol. Aesthet. Creat. Arts* 3, 52–56. doi: 10.1037/a0014835

Tuch, A. N., Presslaber, E. E., Stöcklin, M., Opwis, K., and Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: working towards understanding aesthetic judgments. *Int. J. Hum Comput Stud.* 70, 794–811. doi: 10.1016/j.ijhcs.2012.06.003

Unipark (2017). Questback GmbH (2017). EFS Survey. Köln: Questback GmbH. Available at: http://www.unipark.info (Accessed November 16, 2020).

Valdez, P., and Mehrabian, A. (1994). Effects of color on emotions. *J. Exp. Psychol. Gen.* 123, 394–409. doi: 10.1037//0096-3445.123.4.394

Van Dongen, N. N. N., and Zijlmans, J. (2017). The science of art: the universality of the law of contrast. *Am. J. Psychol.* 130, 283–294. doi: 10.5406/amerjpsyc.130.3.0283

Verhavert, S., Wagemans, J., and Augustin, M. D. (2018). Beauty in the blink of an eye: the time course of aesthetic experiences. *Br. J. Psychol.* 109, 63–84. doi: 10.1111/bjop.12258

Weichselbaum, H., Leder, H., and Ansorge, U. (2018). Implicit and explicit evaluation of visual symmetry as a function of art expertise. *I-Perception* 9:204166951876146. doi: 10.1177/2041669518761464

Winston, A. S., and Cupchik, G. C. (1992). The evaluation of high art and popular art by naive and experienced viewers. *Vis. Arts Res.* 18, 1–14.

Yang, C. C., and Rodriguez, J. J. (1995). "Efficient luminance and saturation processing techniques for bypassing color coordinate transformations." in *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*. 1, 667–672.

Conflict of Interest: 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.

Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs

Claire Reymond

Department of Psychology, Center for General Psychology and Methodology, University of Basel

Jan B. Vornhagen

Department of Computer Science, Aalto University

Matthew Pelowski

Faculty of Psychology and Cognitive Sciences Hub, University of Vienna

Klaus Opwis

Department of Psychology, Center for General Psychology and Methodology, University of Basel

Elisa D. Mekler

Department of Computer Science, Aalto University

Author Note

Claire Reymond https://orcid.org/0000-0002-8988-7263

Jan B. Vornhagen https://orcid.org/0000-0002-8516-7840

Matthew Pelowski https://orcid.org/0000-0001-5563-3727

Klaus Opwis https://orcid.org/0000-0003-0509-8070

Elisa D. Mekler https://orcid.org/0000-0003-0076-6703

The authors thank Beat Vollenwyder for his support with programming the survey on Unipark.

Images Influencing Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs

Abstract

Images are never seen in isolation. Instead, they are perceived within a spatial and temporal tapestry of neighboring images. What impact do other images have on our emotional response toward a particular image? Answers to this basic question have vital implications for a range of fields—especially for visual communication and for art, where resources are invested in arranging images within a visual context. Previous studies have provided mixed results, suggesting that juxtaposed images may lead to contrast or assimilation processes increasing and decreasing our liking of an image. But how specific image features in neighboring images (image's ambiguity or formal similarities between images) modulate our affective interpretation of an image has almost never been explored. In Study 1, we compared the emotion perceived in art photographs ("target" images) when displayed on their own versus when displayed in juxtaposition with negatively or positively valenced nonart ("context") images. Additionally, we analyzed the influence of the artwork's perceived ambiguity. In Study 2, we examined the effect of the perceiver's expertise and the formal similarity between the images on the rated valence of the target image. Our results show that the emotion perceived in the artwork contrasted away from or assimilated toward the valence perceived in the context image depending on which evaluative dimension was activated. Moreover, the influence of negative contextual material on the target image's valence was more pronounced. We conclude by saying that the evaluative dimension is part of the pictorial context that influences the affective interpretation of an image.

Keywords: emotion perceived in art photographs, juxtaposition, context, ambiguity, formal similarity

Running Head: Images influencing images

Images Influenced by Images: How Pictorial Context Affects the Emotional Interpretation of Art Photographs

Introduction

“Whatever the artist may do, however, he cannot avoid showing his surface in the midst of other surfaces of an environment. A picture can only be seen in a context of other non-pictorial surfaces.”
James J. Gibson in the Ecological Approach to the Visual Perception of Pictures, 1978

Images can emotionally move us, intellectually challenge us, or sometimes even change our view of the world. But images are never perceived in isolation. They are rather always experienced embedded in a context. Contextual information such as the physical space in which a work is encountered (e.g., Gartus et al., 2015), the information provided next to the artwork (e.g., Cupchik et al., 1994; Szubielska & Sztorc, 2019), or the judgments of others (e.g., Lauring et al., 2016) influences the aesthetic experience of an image (for an overview on contextual influences on art perception, see Pelowski & Specker, 2020). Specifically, and expanding on the quote by Gibson, wherever we look at images—whether on the street, on a screen, or in a book—they are always accompanied by myriad other visual artifacts: the context in which images are encountered consists of other images. This is especially true for art, where images are most often seen in a gallery or museum and encountered within a specific progression of other images.

The value, and perhaps one of the purposes, of art is to touch us emotionally (Pelowski et al., 2020). At the same time, how images influence the perceived emotional interpretation of neighboring images is a complex phenomenon, involving a multitude of aspects (e.g., influences can be based on formal features of the image like format, color, and style or content-related factors like the image statement and emotional expression). This poses a unique challenge for professionals that deal with the presentation of (art) images. Curators and gallery owners spend time and effort arranging images next to each other for exhibitions and online museum tours (e.g., Flacke, 2016), and designers put a lot of thought into how to present images on double-page spreads in catalogues, books, and newspapers (e.g. Samara, 2007). At the same time, professionals who study and work intensively with the presentation of images are (implicitly) aware of interaction effects between images, and they compose image series in order to intensify, contradict, or change their evaluation and message (Ganz & Thürlemann, 2010; Hofmann, 1985; Reymond, 2013). The knowledge gained from our study thus supports curators in their aim to shape encounters with art in exhibition contexts, and it may enable designers to employ the interplay of images in visual design more consciously to afford an intended effect. This will in turn impact the emotions experienced as part of our interaction with (art) images. However, there is a lack of empirical studies examining the effect of pictorial context and the modulating effect of specific image features on the emotion perceived in an image. Even more, existing studies have led to contradictory results, which indicates a need for more controlled and targeted research on images' impacts on other images and the resulting affective interpretations. This is the aim of the present paper.

Pictorial Context and its Impact on Perceived Emotions

In this paper, *pictorial context* refers to the spatial and temporal environment of a stimulus consisting of images shown next to each other or in a sequence (Cohn, 2013, 2015). Contextual information—often presented as images—has been shown to influence the perception of emotions in faces (Barrett et al., 2011): for example, an expression of disgust paired with a muscled body was interpreted as a proud face, which shows that the emotions perceived in images of faces are interpreted according to the context in which they are encountered (Aviezer et al., 2008). At the same time, a photograph of a fearful face is evaluated as more fearful surrounded by an image depicting a threatening situation than in a happy or neutral situation (Righart & de Gelder, 2008). In film, the influential power of sequential images is known as the Kuleshov effect (Kuleshov, 1974;

Mobbs et al., 2006): the affective interpretation of identical moving images of a face is changed by the sequence viewed before it containing highly emotional material. Recent studies have replicated the Kuleshov effect by demonstrating that the perceived valence and arousal of moving images showing neutral faces depended on previously viewed scenes (Calbi et al., 2017), and faces seen in a negative or positive context elicited higher valence and arousal reactions than faces in neutral contexts (Mobbs et al., 2006). In a study on the perception of a work of art, the facial expressions of two depicted figures were judged on the basis of the context in which they were seen: manipulating the position of a fearful figure affected the perceived emotional interpretation of another figure in the same image (Marian & Shimamura, 2012).

Similarly, images accompanying textual news reports in mass media have been shown to systematically influence the (emotional) understanding of the written message (Price et al., 1997). Images can serve as frames for interpreting neighboring text, since they use various rhetorical means—metaphors, illustrations, symbols—that graphically capture the essence of an event (Rodriguez & Dimitrova, 2011, p. 51). The emotional valence of a picture placed next to a text influences the subsequent cognitive processing of information such “that reactions to featured photographs shift the primarily text-based perceptions and evaluation of issues in the direction suggested by the photographs” (Gibson & Zillmann, 2000, p. 355).

Perceptual Processes Underlying the Impact of the Pictorial Context

The influence of the pictorial context on the perception of an image can be explained as a categorization of neutral faces according to the emotions attributed to the context (Calbi et al., 2017) or as a visual frame used to interpret the meaning of text (Rodriguez & Dimitrova, 2011). Specifically, two processes have been proposed:

First, images that are seen after very favorably rated images are perceived as less favorable. In that case, the perceived image (target image) forms a *contrast* to the image that is seen in the context (context image). Paintings from Goya's tapestry period were liked more when they were seen after works from his dark period (Dolese et al., 2005); beauty ratings for a set of moderately beautiful photographs of buildings were higher after viewing a set of less beautiful pictures (Tousignant & Bodner, 2014); and abstract paintings were rated more beautiful when paired with less beautiful paintings (Tousignant & Bodner, 2018).

This effect of stimuli contrasting *away* from contextual stimuli was described by Fechner's principle of aesthetic contrast (Fechner, 1898 in Allesch, 2018). Fechner's principle describes a hedonic contrast. It distinguishes between positive hedonic contrasts, when a stimulus is rated better after seeing a contextual stimulus, and negative hedonic contrasts, when a stimulus is evaluated as worse after perceiving a contextual stimulus (Dolese et al., 2005). Parducci's range-frequency model (1965) provided an explanation for Fechner's principle of aesthetic contrast and aligns it with some general psychological heuristics such as the anchoring effect (for an overview, see Furnham & Boo, 2011). Parducci (1965) proposed that a stimulus will be rated on a fictitious rating scale relative to other recently rated stimuli and that both the range (distance between the most positive and most negative stimuli) and the distribution of recently rated stimuli has an influence on how a stimulus is rated. That is, an extremely positive or negative context stimulus will push the target's rating down or up and so provoke contrast effects. Also, if the other stimuli are evenly distributed (vs. not evenly distributed), a neutral stimulus will be placed closer (vs. more far away) to the middle of the rating scale.

Second, an image may assimilate *toward* an image seen before or next to it. Images that were preceded by highly negative IAPS (International Affective Picture System; Lang et al., 1997) context images assimilated toward the negative context images (Mullennix et al., 2018). When rating the pleasantness of an artwork, contrast effects occurred when the artwork was presented next to contextual stimuli that were formally similar to the target but clearly aesthetically inferior to it. Assimilation effects were registered when the contextual stimulus was aesthetically similar to the target stimulus (Arielli, 2012). Assimilation effects were also found based on participants' tendency to repeat the previous response (Chang et al., 2017; Kondo et al., 2012). Pegors et al. (2015) showed an assimilation effect to previous evaluations, but at the same time the stimulus qualities that participants had viewed in the preceding trial had a contrasting effect on the judgment of the current stimulus.

Several models have been proposed to describe the effects of assimilation and contrast (Förster et al., 2008; Schwarz & Bless, 1991). One of them is Mussweiler's (2003) selective accessibility model. Mussweiler's model is based on a three-stage process: the selection of a context stimulus to which the target stimulus will be compared, the comparison between the two, and the evaluation of the comparison. Applied to images, this suggests that perceived similarity between a target image and context image facilitates assimilation, that is, the target image is rated as *more similar* to the context image. Conversely, perceived dissimilarity between two images facilitates contrasting evaluations, in which the target image is perceived as *more different* from the context image.

Some studies have shown that negative stimuli have a more pronounced effect on subsequent ratings of stimuli than positive ones. In a study by Calbi et al. (2017), presenting happy or fearful faces before neutral faces had a clear effect on the assessment of the neutral faces, but only in the fear condition. Similarly, Mullennix et al. (2018) only found an assimilation effect when target images were shown next to negative image material and not when they were paired with positive images. This effect may be explained with the motivated attention theory (Bradley et al., 2003; Lang et al., 2013; Schupp et al., 2004). Cues indicative of danger and fear evoke response facilitation compared to neutral stimuli and motivate a defensive attitude, which become manifest in avoidance and heightened vigilance (Bradley et al., 2012). The comparably stronger impact of negative image material is consistent with the negativity bias, which explains this effect as a function of evolutionary adaptation (Baumeister et al., 2001; Vaish et al., 2008).

Curators and designers have an interest in presenting artworks in a way that supports, challenges, or contradicts the images' inherent meanings and in evoking genuine emotional reactions to them. By using formal analogies or contrasting contents, they shape interactions between images, creating possibilities of influence between pictures that stand side by side. But this knowledge largely remains tacit, based on curators' and designers' many years of experience and practical work with images. A better understanding of which processes underlying the effects of pictorial context are activated under which circumstances may therefore help inform curatorial decisions and facilitate teaching novice image practitioners.

Present Studies

We investigated two distinct factors that might account for how pictorial context shapes viewers' emotional attributions in images. In Study 1, we examined the role of ambiguity on perceived emotion in juxtaposed images. In Study 2, we investigated formal similarity between images as a potential moderator for how pictorial context impacts viewers' evaluation of art photographs. We did this by measuring the emotion perceived in a target image when participants were confronted with an emotionally negative or emotionally positive context image compared to when the target image was viewed on its own.

Study 1: The Effect of Ambiguity on Valence Ratings in Juxtaposed Presentation

Ambiguity refers to when multiple meanings are attributed to an object and the meaning varies depending on the information, context, and interaction between an observer and an object (Gaver et al., 2003). In images, cognitive ambiguity is a visual experience that elicits multiple interpretations (Jakesch et al., 2013). Since works of art exhibit semantic instability (Jakesch & Leder, 2009; Muth et al., 2015), they are predisposed to be affected by contextual influences. According to Herr et al. (1983), contextual information can be activated as a prime. If the prime is moderately extreme, “it is the ambiguity of the target that determines whether assimilation or contrast effects emerge” (p. 334).

The aim of Study 1 was therefore to investigate the extent to which image ambiguity moderates the influence of the pictorial context—in the form of negatively vs. positively valenced context images shown in juxtaposition—on the emotional attribution of a target image.

Method

Participants. Study 1 included 106 participants recruited via Prolific (prolific.co; mean age = 26.3, *SD* = 7.3, 38 female, 66 male, two nonbinary). This sample size was informed by a simulation study using the BEST package (Kruschke & Meredith, 2021). We used the default priors to estimate a required sample needed for excluding artworks with unreliable ambiguity ratings (i.e, ambiguity ratings with an *SD* > 1 on a 7-point Likert-type scale). For Study 1, this resulted in a recommended sample size of $n \geq 54$ per group for a between-participant design.

All participants had normal or corrected-to-normal vision as well as normal color vision. Participants received monetary compensation (about US$2) for enrolling in the study. Both studies were approved by the ethics committee of the University of (anonymized for peer review).

Stimuli. Photographs depicting landscapes and scenes were used as stimuli for both studies. The focus of our study was to examine the affective interpretation of images seen with negatively or positively valenced context images. However, we did not include context images that were likely to elicit strong emotions (e.g., mutilated bodies or dangerous animals), because we wanted to investigate a more realistic situation of images that could be presented side by side (for studies using highly emotional stimuli, see Mullennix et al., 2018). We therefore excluded images showing humans and animals.

The stimuli in Study 1 consisted of 20 target images and 20 context images. The target images were horizontal format high-quality (150 dpi) digital color photographs of fine art by contemporary Western artists (see "Art Photographs" for a full list of the artworks, artists, and links to the retrieved photographs in the repository: https://osf.io/ptfqe/?view_only=f8be94e99d074065b74d1e1c8b34e5f3.). The images were downloaded from the artist's own website, from museum's websites, or from auction houses. As such, the images were expected to represent the most faithful reproductions of the original photograph's contrast and color (Reymond et al., 2020). Context images were selected from the OASIS Image Set (Open Affective Standardized Image Set; Kurdi et al., 2017) and consisted of horizontal format (nonart) color photographs. We used context images that were prerated as either having a negative (mean valence = 1–2.5 / 7) or a positive (mean valence = 5.5–7 / 7) valence. Prerated arousal values were included in the analysis as a covariable.

Participants were not informed about the provenance of either set of photographs (i.e., they were not informed whether the images were categorized as art or not). We did this for two reasons: first, to obtain more pronounced valence ratings, as it is known that images classified as nonart receive more extreme valence ratings (Gerger et al., 2014; Leder et al., 2014; Pelowski, et al., 2017a). Second, and more importantly, we did not label them as art so as to avoid establishing two categories of stimuli, art images and nonart images, which could introduce potential confounding factors (Dolese et al., 2005; Zellner et al., 2003, 2009).

The target images were displayed next to context images, as it has been shown that the effect is more pronounced if the images are shown juxtaposed rather than sequentially (Khaw & Freedberg, 2018; Tousignant & Bodner, 2018). Ten target images were shown paired with 10 negatively valenced context images (OASIS, Kurdi et al., 2017), and another 10 fine-art photographs were paired with 10 positively valenced context images.

Procedure. The experiments were designed using Unipark software (Unipark, 2017). After providing informed consent, participants were informed that they would view images and be asked to rate them. Study 1 consisted of three blocks presented in the same order for all participants.

Block 1 contained two subblocks (Block 1a [ambiguity rating] and Block 1b [valence rating]). The subblock order was counterbalanced across subjects. In both subblocks, participants were presented each of the 20 target images alone. All images were shown centered on a white background for an indefinite amount of time, and participants could take as long as they wished to look at and rate the image (Arielli, 2012; Dolese et al., 2005; Mullennix et al., 2018). After answering the question presented underneath the image, participants clicked “Continue” at the bottom of the page to see the next image.

In Block 2, participants were asked to provide demographic information (age, gender) and to indicate whether they were professionally involved in assessing or creating images (i.e., as designers, curators, photographers, etc.). We asked participants to state their current affect (PANAS; Thompson, 2007) as well as their tolerance to ambiguity (MSTAT-II; McLain, 2009). Block 2 was also defined to distract participants from their ratings in Block 1.

Block 3 began with the instruction that now two images (a target image and a context image) would be presented side by side and that one of the two images (the target image) should be evaluated in terms of perceived emotion. Following this, participants saw 10 target images paired with 10 formally similar OASIS images (Kurdi et al., 2017) with negative valence and 10 target images paired with formally similar OASIS images with positive valence. Participants could view the target and context images for as long as they wanted. All the image pairs were shown one after another in a randomized order. To make sure that participants rated the valence of the target image and not the context image, a black bar (4 px) was placed under the target image. Whether the target or the context image was placed on the left or on the right was fully randomized.

Finally, participants were asked to indicate if they had recognized any of the photographs (to control for effects of familiarity) and to state if they had answered the questions conscientiously. On average, it took participants about 15 minutes to complete the study.

Measures. We collected measurements on five different scales during three phases of our first study. In Block 1 and 3, ratings were made using a slider displayed beneath the image(s). In Block 1a participants were asked to indicate how ambiguous (0 = "not at all," 100 = "very much") they perceived the displayed image to be. Ambiguity was described as "to what extent an image allows for multiple interpretations and meanings" (Jakesch et al., 2013; Muth et al., 2015). In Block 1b, participants were asked to state the emotion they perceived in the photograph (0 = "very negative," 100 = "very positive"), regardless of the emotion it aroused in them (Gabrielsson, 2001; Kallinen & Ravaja, 2006).

In Block 2, we used the PANAS short version (Positive and Negative Affect Schedule; Thompson, 2007) to assess positive and negative affect, since the perception of art is influenced by the perceiver's affective state (Konečni & Sargent-Pollock, 1977; Leder et al., 2004). Eight items ("ambiguous stimuli in general," "insoluble/illogical/irreducible/internally inconsistent stimuli") of the MSTAT-II questionnaire (McLain, 2009) were used to measure participants' tolerance toward ambiguous stimuli.

During the last phase of the study (Block 3), participants rated the emotion they perceived in a target image next to a context image, regardless of the emotion it aroused in them on a scale from 0 ("very negative") to 100 ("very positive"). For all the ratings on images, participants were asked to note their first impression in order to assess their initial response to the image.

Results and Discussion Study 1

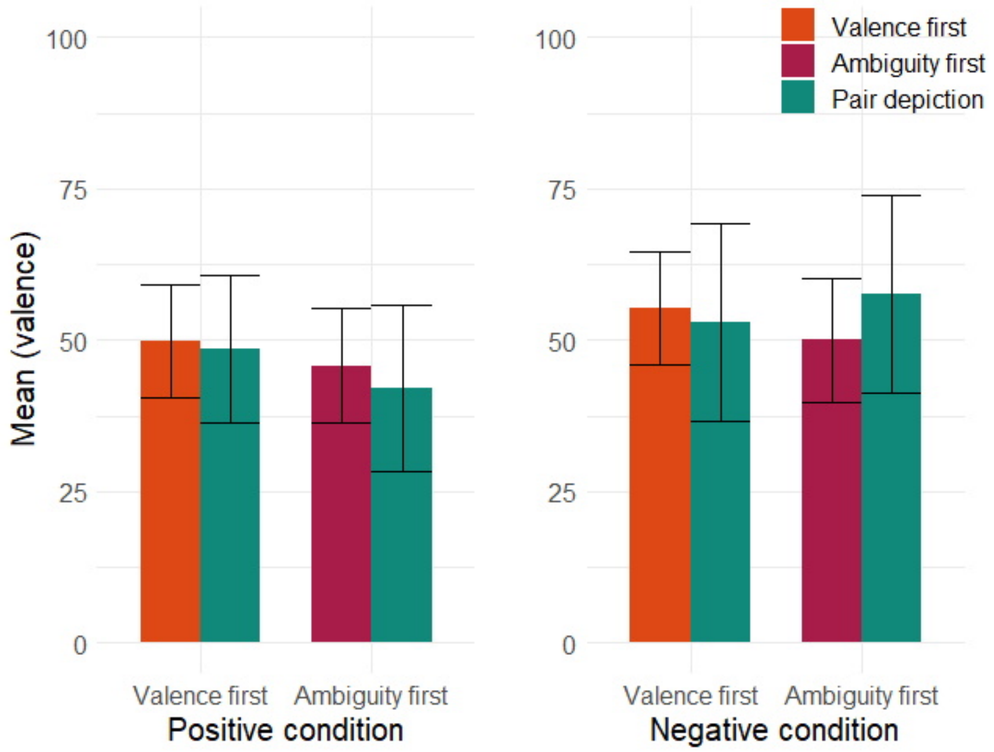
The complete analyses and data are available at: https://osf.io/ptfqe/?view_only=f8be94e99d074065b74d1e1c8b34e5f3. For the analysis, we utilized R 4.1 (R Core Team, 2021) and additional packages for data handling (Comtois, 2021; Pedersen, 2020; Wickham et al., 2019) and Bayesian analysis (Bååth, 2015; Kruschke & Meredith, 2021). We opted for a Bayesian approach due to the absence of clear confirmatory hypotheses. We opted for computing highest density intervals (HDIs) to estimate the most probable effect sizes for the effects of interest. Compared to a frequentist approach, which usually computes the probability of the data given the null hypothesis, $P(D|H)$, the Bayesian approach computes the probability of the hypothesis given the observed data and prior beliefs, $P(H|D)$ (Dienes, 2008). We therefore used Bayesian inferences to estimate the probability of certain effect sizes given our data.

Descriptive Statistics. The descriptive analysis of the short version of the PANAS (Positive and Negative Assessment Scale; Thompson, 2007) showed that participants scored generally low on negative affect (M = 1.57, SD = 0.93) and medium on neutral and positive affect (M = 3.28, SD = 1.13). Similarly, participants indicated a moderate tolerance for ambiguous stimuli in general (MSTAT-II G, M = 3.95, SD = 0.7; MSTAT-II I, M = 3.36, SD = 1.25). None of the participants indicated they knew any of the presented stimuli; two participants indicated that they were unsure if they knew a stimulus, but they did not correctly guess what the stimuli were.

Main Analysis. To explore our first question—how the ambiguity of the target image moderates the effect of the context image—we compared the ambiguity ratings with the valence ratings. We also examined whether the order in which the images were rated (i.e., ambiguity first versus valence first) affected participants' ratings. No meaningful correlation was found between the ambiguity of the target image and the changes in valence from the presentation of single images to pairwise presentation (positive condition, r =.026; negative condition, r = −.014).

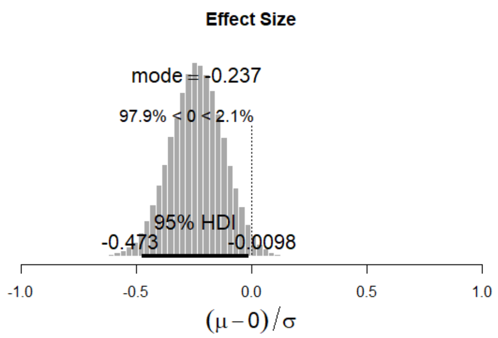
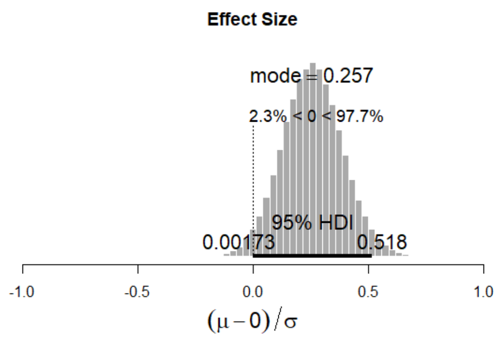
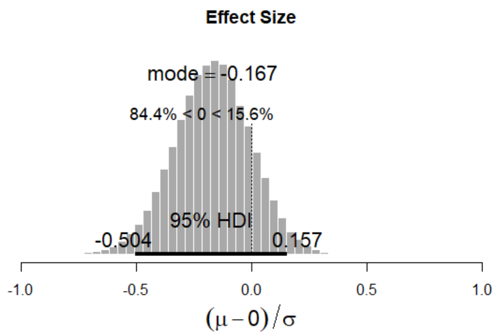
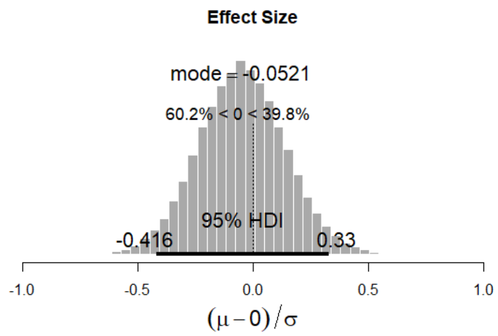
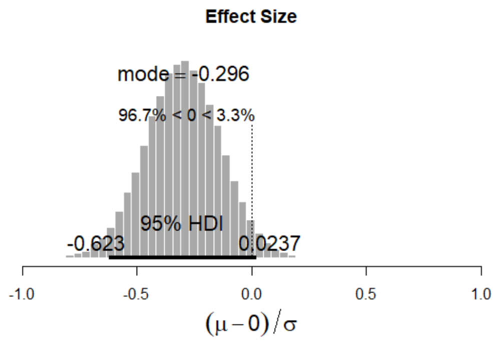
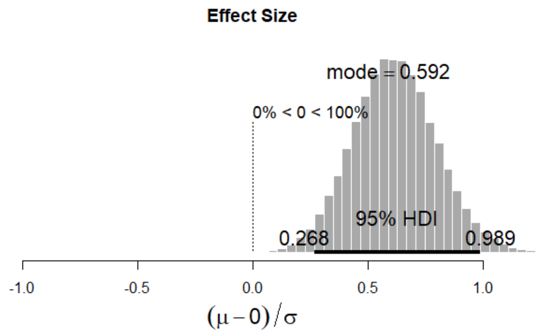
However, as shown in Figure 1, the order in which participants rated ambiguity and valence resulted in the artworks being rated differently in the paired condition. When ambiguity was rated first, participants changed their valence ratings more strongly from the solo presentation to the pairwise presentation. Furthermore, when the images were judged first for their ambiguity and then for their valence, the effect of juxtaposition produced contrast effects for both the positive and the negative conditions. When the valence of the images was judged first, a slight tendency to an assimilation effect was observed in the negative condition. The overall effect was more pronounced when the target images were seen next to negative image material than next to positively valenced images (Fig. 1).

Figure 1: Mean ratings and standard deviation for valence over all stimuli when presented alone in the valence-first condition (orange) and the ambiguity-first condition (red) and when presented paired with a context picture (green).



To answer our second question—Does the presence of a juxtaposed context image influence the emotion perceived in a target picture?—we compared the valence ratings in the single-image condition to the paired condition. First, we applied Bayesian analysis to estimate the size of the difference between the rating of a target stimulus seen on its own and the rating when it was seen with a context picture (Table 1). We used a flat prior and computed 95% HDIs to estimate the extent to which the valence ratings of the target images changed when they were presented with a context image. The estimated effect size of the difference was computed by subtracting the rating of the single image from the paired rating, so a negative effect size indicates a more positive rating in the single condition, and positive effect size indicates a more positive rating in the paired condition.

Table 1: Shown are the 95% HDIs for the estimated effect sizes of the differences between the ratings of the paired and the solo presentations of the images. The effect sizes are categorized by the valence of the context image and if valence was rated first (valence first) or second (ambiguity first) during the first block. A negative effect size indicates a more positive rating in the solo condition.

Positively valenced context image	Negatively valenced context image
Overall	
	
Valence first	
	
Ambiguity first	
	

Compared to when displayed individually, the target images were rated more negatively when paired with a positively valenced context image (mode = −0.24, HDI_{low} = −0.47, HDI_{high} = −0.01) and more positively when presented with a negatively valenced context image (mode = 0.26; HDI_{low} < 0.01, HDI_{high} = 0.52). When taking the rating order (ambiguity first vs. valence first) into account, the estimations of the effect sizes change for valence first: there was a slightly negative effect size (mode = −0.05, HDI_{low} = −0.42, HDI_{high} = 0.33) when the target images were presented with negatively valenced context images. When the context image was positively valenced, the target images were rated slightly less positive in the paired condition compared to the single-image condition (mode = −0.17, HDI_{low} = −0.5, HDI_{high} = 0.16). In the ambiguity-first group, the target images were rated more negative when paired with a positively valenced context image (mode = −0.3., HDI_{low} = −0.62, HDI_{high} = 0.02). The target images were also rated more positive when paired with a negatively valenced context image (mode = 0.6, HDI_{low}=0.27, HDI_{high} = 0.99).

In Study 1, we examined image ambiguity as a moderating aspect in the juxtaposition effect. In Study 2, we investigate the effect of formal similarity on perceived valence in the paired condition.

Study 2: Effect of Formal Similarity on Valence Ratings in Paired Images

What underlies contrast and assimilation effects is the premise of similarity in the rated stimuli. In past studies, similarity has been considered in terms of category membership (Dolese et al., 2005; Mussweiler & Bodenhausen, 2002; Rota & Zellner, 2007) or how extreme the standard is (Herr et al., 1983). But contrast effects have also been observed irrespective of whether the context images were assigned to a same, a similar, or a different category (Tousignant & Bodner, 2014, 2018). These results suggest that similarity may be an elastic concept in the evaluation of images.

Categorical similarity evaluates the correspondence of conceptual factors, which are based on knowledge or expertise and processed top-down (Leder et al., 2004; Pelowski, et al., 2017b). In contrast, formal similarity describes the similarity of basic image features, which are assessed in a bottom-up manner at the beginning of the image evaluation process (Pelowski et al. 2017b). Formal similarity encompasses diverse image features, such as color, orientation, size, and motion, and can be perceived in time or space (Arnheim, 1957). According to Mussweiler (2003), comparisons between context and target stimuli are based on a rapid search for salient similarities. It is therefore to be expected that bottom-up processed image features such as color and shape influence the comparison process on a basic level.

Crucially, image expertise affects the perceived similarity between stimuli and therefore influences context effects. Because experts use as their basic-level categories what are the subordinate levels for novices, they recognize similarities that novices do not see. In a study assessing the attractivity of flowers, this resulted in a hedonic contrast for laypeople but not for experts (Rota & Zellner, 2007). More specifically, unlike laypeople, design experts are expected to recognize formal similarities, since they are trained to judge images not only in terms of their content but also in terms of their formal properties and similarities (Arielli, 2012; Hofmann, 1985).

In our second study, we analyzed the role of formal similarity between target and context images as a moderating factor for the emotion perceived in target images that are presented juxtaposed to a context image. We assumed that formal similarity between the target and context image would enhance the influence of the context image on the target image. This effect was expected to be strongest between a target image and a formally similar context image with negative emotional valence. We compared ratings by laypeople and design experts.

Method

Participants. We used the expected valence means of the target pictures as priors to estimate how many people were needed to reliably detect a deviation of 1 from the single image ratings on a 7-point Likert scale. This resulted in a recommended sample of $n \geq 54$ per group.

Consequently, 122 participants were recruited via Prolific (prolific.co). The mean age was 33.7 (*SD* = 10.39), and there were 90 women and 32 men. In addition, we recruited 56 design experts (design students and lecturers from the University of (anonymized for peer review)) (mean age = 28.95, *SD* = 9.78, 36 female, 19 male, one nonbinary). This resulted in a final sample size of *N* = 178, of which 56 people identified as experts (46 of these experts were from the design-experts sample and 10 were from the Prolific sample). The experts from each sample did not identifiably differ from each other and were consequently treated as a singular expert group (see Study2_Analysis in the repository). All participants had normal or corrected-to-normal vision as well as normal color vision. Participants recruited via Prolific received monetary compensation (about US$2) for completing the study.

Of the Prolific sample, one participant indicated that they had seen one of the stimuli before, but they did not identify it correctly. Of the design experts, 14 (25%) indicated that they recognized artworks, and nine of them correctly recognized a few of the works. Since there were less than 10 persons in the total sample who recognized one or more of the art images correctly, we did not pursue this further.

Stimuli. The same fine art photographs and context images as in Study 1 were shown in pairs. To obtain formal similarity between the target and context images, the selection process of the images was conducted in two phases: first, context images with required valence values were selected from the "scenes" category of the OASIS (Kurdi et al., 2017) dataset. In a second phase, art photographs were chosen that resembled the selected context images in motif, color, and form. The selection process was done by the first author in collaboration with a lecturer in design at the University of (anonymized for peer review) and in accordance with the other authors. The formal similarity between the images was pre-assessed by five individuals not involved in the study.

Four experimental pairing conditions were defined. In the negative formal similarity condition (A), a fine art photograph (target image) was shown with a negatively valenced OASIS image (context image) that was formally similar (Fig. 2). In the negative no formal similarity condition (B), a fine art photograph was shown with a negatively valenced context image that was not formally similar (Fig. 3). In the positive formal similarity condition (C), an art photograph was presented with a positively valenced context image that was formally similar. In the positive no formal similarity condition (D), one art photograph was presented with a positively valenced context image that was not formally similar. Each of the 20 fine art photographs was assigned either to the positive-valence condition or the negative-valence condition (10 positive and 10 negative). Within these conditions, the same target image was paired once with a formally similar context image and once with a formally dissimilar context image (See Figs. 2 and 3). Thus, 20 target images (all fine art photographs) and 40 context images (all OASIS images) were used (see "Image Pairs" in the repository for a detailed overview). All the target pictures had been rated regarding their valence and ambiguity in Study 1.

Figure 2: Example of pairing condition A in Study 2. Target image on the left: Thomas Keller, *Ohne Titel*, from the series *Häuser – Where Distance Lives*, 2004–2010, www.thkeller.com; paired with a formally similar context image with negative valence on the right; OASIS Image Set, Flood 3.

Figure 3: Example of pairing condition B in Study 2. Target image on the left: Thomas Keller, *Ohne Titel*, from the series *Häuser – Where Distance Lives*, 2004–2010, www.thkeller.com; paired with a negatively valenced context image that is not formally similar on the right; OASIS Image Set, Explosion 2.

Procedure. Study 2 consisted of three blocks presented in the same order for all participants. In Block 1, participants were presented with each of the 20 target images paired in four different conditions (A = negative formally similar condition, B = negative not formally similar condition, C = positive formally similar condition, D = positive not formally similar condition). After the first block was completed, participants were asked in Block 2 to provide demographic information (age, gender) and to rate their current affect (PANAS; Thompson, 2007). Participants were also asked to indicate if they had recognized any of the images they saw in the study and whether they were professionally involved in assessing or creating images.

In Block 3, participants were presented with 20 pairs of a target and a context image displayed side by side. To make sure that participants rated the valence of the target image and not the context image, a black bar (4 px) was placed under the target image. Whether the target or the context image was placed on the right or on the left was fully randomized. All the images shown in this study were presented in a randomized order.

To conclude the experiment, participants were then asked to indicate if they had recognized any of the fine-art photographs (to control for effects of familiarity) and to state if they had answered the questions conscientiously.

Measures. We collected measurements on three different scales during the three phases of Study 2. In Block 1, participants were asked to indicate how formally similar (0 = "not at all," 100 = "very much") the two images were on a sliding scale displayed beneath the images. Formal similarity was described as how much the two photographs looked formally similar and that this included "qualities of composition such as color, shape, form and line" (Cupchik et al., 1992).

During Block 2, participants rated their current affective state using the 10 items from the PANAS (Thompson, 2007). In Block 3, participants were asked to rate the emotion they perceived in the target photograph, regardless of the emotion it aroused in them on a sliding scale from 0 ("very negative") to 100 ("very positive"). When rating formal similarity and perceived emotion, participants were instructed to note their first impression of the image.

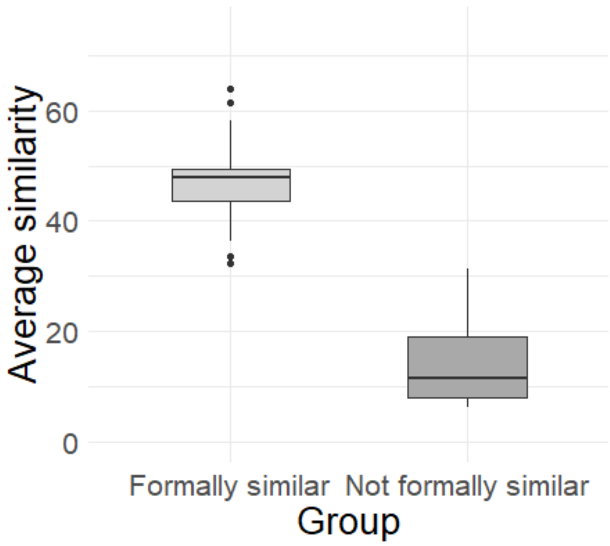
Results and Discussion Study 2

As in Study 1, the complete analyses and data can be found at: https://osf.io/ptfqe/?view_only=f8be94e99d074065b74d1e1c8b34e5f3. For analysis, we utilized R 4.1 (R Core Team, 2021) and additional packages for data handling (Comtois, 2021; Pedersen, 2020; Wickham et al., 2019) and Bayesian analysis (Bååth, 2015; Kruschke & Meredith, 2021).

Descriptive Statistics. Test-retest reliability, as measured by the repeated images in Block 1 and Block 3, showed that all the image ratings had good reliability (total, r =.76, n = 244). Hence, all the data were retained for analysis. Results from the short version of the PANAS (Thompson, 2007) indicated that the participants—comparable to Study 1—scored generally low on negative affect (M = 1.58, SD = 0.92) and medium on neutral or positive affect (M = 3.35, SD = 1.1).

Main Analysis. To answer our research question for Study 2—Does formal similarity moderate the juxtaposition effect?—we first examined the subjective formality ratings. We compared the similarity ratings of the design experts (n = 56) and laypeople (n = 122). The design experts rated formal similarity comparably to the laypeople: not formally similar condition, M = 13, *SD* = 16.7 vs. *M* = 14, *SD* = 18.8; formally similar condition, *M* = 47.9, *SD* = 25.4 vs. *M* = 46.9, *SD* = 27.3. Thus, we aggregated the two groups as one in the following analysis. As depicted in Figure 4, the formally similar pairs—while perceived as clearly more similar than the pairs that were not formally similar—were rated to be moderately similar, scoring between 40 and 60 on formal similarity.

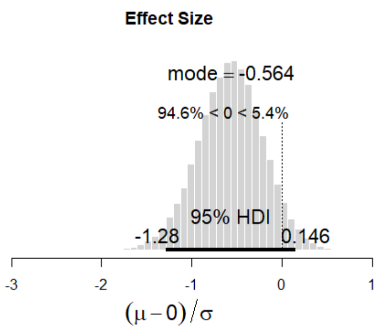
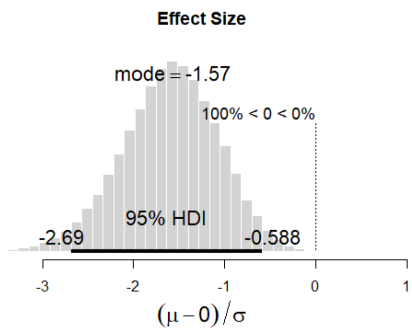
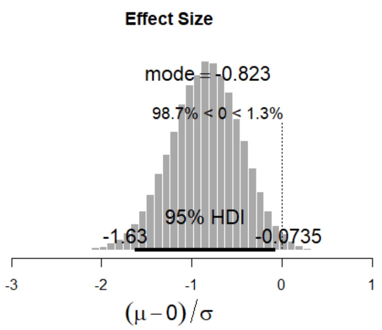
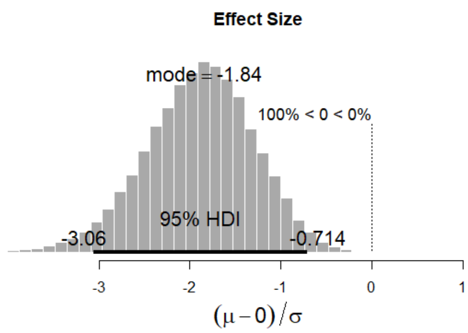
Figure 4: Average rating of the formal similarity of the fine art photographs and their assigned pair. Separated into formal similarity and no formal similarity.



We then compared the valence measurements for the single-presentation condition from Study 1 with the valence measures in the juxtaposed condition in Study 2 for each of the four conditions (*n* = 10): A = negative formally similar condition, B = negative not formally similar condition, C = positive formally similar condition, D = positive not formally similar condition. Since the effect size of the difference was computed by subtracting the valence scores in Study 1 from the valence scores in Study 2, a negative effect size indicates that the artworks were rated more positive when presented on their own than in the juxtaposed condition (see Table 2). A clear effect can be observed in all four conditions. Similarly to Study 1, the effect was more pronounced when the target image was paired with a negative context image (formally similar condition, mode

= −1.57, HDI_{low} = −2.69, HDI_{high} = −0.59; not formally similar condition, mode = −1.84, HDI_{low} = −3.06, HDI_{high} = −0.71) than when the artwork was paired with a positively valenced context image (formally similar condition, mode = −0.56, HDI_{low} = −1.28, HDI_{high} = 0.15; not formally similar condition, mode = −0.82, HDI_{low} = −1.63, HDI_{high} = −0.07). All the effect sizes correspond to a large effect according to Cohen (1992). As can be seen in Table 2, formal similarity between the target and context image did not have a substantial effect. On the contrary, the effect was slightly more pronounced in the conditions where there was not any evident formal similarity between the images (the differences between the individual artworks can be found in the repository; Study2_Analysis).

Table 2: The 95% HDIs for the differences between the paired ratings in Study 2 and the solo valence ratings from Study 1 of the same stimulus. Categorized according to the valence and the formal similarity of the context image. A negative effect size indicates a more positive rating in the solo condition.

<table>
<tr><td></td><td>Positively valenced context image</td><td>Negatively valenced context image</td></tr>
<tr><td>Formally similar</td><td>
</td><td>
</td></tr>
<tr><td>Not formally similar</td><td>
</td><td>
</td></tr>
</table>

Next, we compared the design experts (n = 56) and laypeople (n = 122) regarding their valence ratings. Comparably to the similarity ratings, experts and laypeople rated perceived emotions similarly: positive, M = 45.1, SD = 23.9 vs. M = 44.6, SD = 24.3; negative, M = 42.8, SD = 23.7 vs. M = 43.2, SD = 23.3. Moreover, the differences in the affect scores (PANAS; Thompson, 2007) between the experts (positive affect, M = 3.01, SD = 1.14; negative affect, M = 1.62, SD = 0.95) and laypeople (positive affect, M = 3.51, SD = 1.03; negative affect, M = 1.56, SD = 0.91) was negligible. As the valence ratings and the affect scores were comparable between experts and laypeople, we did not pursue analysis of expertise effects further. (For a more detailed description of the two groups, please refer to the supplementary material.)

Finally, we explored whether the juxtaposition effect could be tied to the arousal score of the context image. We used a Pearson correlation to investigate the connection between the valence scores in the paired condition with the prerated arousal scores of the OASIS (Kurdi et al., 2017) pictures (see the repository). However, due to only having 10 pairs per condition, the estimations of potential correlations were rather wide, which hindered determining whether such effects were present. While art photographs paired with negatively valenced context pictures slightly skewed toward negative correlations (median $r = -.15$, $HDI_{low} = -.57$, $HDI_{high} = .30$), art photographs paired with positively valenced context pictures skewed toward positive correlations (median r = .25, $HDI_{low} = -.18$, $HDI_{high} = .64$)

General Discussion

The perhaps essential value of art is to touch us emotionally (Pelowski et al., 2020). When we encounter art, however, we rarely perceive art images in isolation. This in turn has an influence on the affective interpretation we attribute to a single piece of art. We conducted a study that reflected a common situation for encountering art: art photographs are often embedded in a pictorial context. Our focus was to manipulate the valence of neighboring images (using prescored negatively or positively valenced context images) and the formal similarity (formally similar or not formally similar) between the target and context images so as to examine the effect of these image features on the emotional attribution of the art image. We also considered the ambiguity of the artworks and the expertise of the viewers.

Juxtaposition Effects on the Affective Interpretation of Art Photographs

In our studies, we clearly showed that neighboring images influence the affective interpretation of art photographs. These results reflect earlier findings that emotionally loaded image material affects the interpretation of faces (e.g., Barrett et al., 2011; Righart & de Gelder, 2008) and text (e.g., Price et al., 1997; Rodriguez & Dimitrova, 2011). But more importantly, our results extend the findings of previous studies that showed an influence of neighboring images on the liking and aesthetic evaluation of images (e.g., Arielli, 2012; Mullennix et al., 2018; Tousignant & Bodner, 2014).

We demonstrated—in line with previous studies (e.g., Calbi et al., 2017)—that negative pictures exert a stronger influence on neighboring images than positive ones do. Although our results exhibited less pronounced effects compared to studies that used unambiguously negative images (Mullennix et al., 2018), our results show that the perception of images and their emotional attribution can be influenced by other images—even if they depict merely subtly negative material. The reason why negative pictures tend to have a stronger impact on neighboring images has been explained as a general negativity bias (Baumeister et al., 2001; Rozin & Royzman, 2001). However, in contrast to Calbi et al. (2017) and Mulllennix et al. (2018), we also found a negative trend when target images were presented alongside context images with positive valence. This allows us to explain the generally stronger influence of negative contextual material on the target image in terms of negativity bias, but the positive condition in our studies shows that we are actually dealing with a contrast effect and not just a general negativity bias. If this effect were merely due to a negativity bias, the ratings of the images in the positive conditions would have remained the same when the target image was presented alone and when it was presented in juxtaposition with the context image. In our studies, however, we found that target images were rated more negative when viewed juxtaposed with positive images than when viewed alone.

Contrast and Assimilation Prompted by Rating Dimension

Our results revealed contrast and assimilation effects. In our first study, we found an overall contrast effect: images shown with a negatively valenced context image were perceived to be more positive than if they were seen alone, whereas images shown with a positive context image were rated more negative than when rated alone. When we separated the ratings into valence-first and ambiguity-first conditions, we discovered a slight tendency for an assimilation effect in the negative condition and a contrast effect in the positive condition when the valence was rated first. This pattern was also found in our second study, where it revealed even larger effects.

Following the selective accessibility model (Mussweiler, 2003), which predicts assimilation when the comparison reveals that the target and context are similar and contrast when they are dissimilar, we assume that in the valence-first condition, the target images were rated as similar to the context images in the negative condition (thus producing assimilation), whereas they were perceived as dissimilar in the positive condition (thus provoking a contrast effect). Accordingly, the target images in the ambiguity-first condition were rated as dissimilar to the context images, thus producing contrast effects in both conditions. Mussweiler (2003) points out that the process of comparison is not always conscious. Comparisons between context and target images may occur spontaneously or even subliminally, and standard features may be “identified, retrieved, or constructed on the spot” (p. 480). Similarly, Higgins and Chairs (1980) proposed that if the recently activated dimension “is applicable to the stimulus (i.e., there is a sufficient match between the features of the construct and the features of the stimulus), then it will be used to encode or characterize the stimulus" (Higgins & Chaires, 1980, p. 351). We therefore argue that the different effects in our first study were possibly due to the order of the rating dimensions and the target knowledge that was thereby activated. In that sense, the rating dimension can be seen, in addition to the neighboring image, as part of the pictorial context in which the target images were assessed, so it too influences the emotional evaluation of the images. Activating ambiguity first seems to have influenced participants to perceive the target and context images as dissimilar in general, whereas activating valence first provoked similarity and dissimilarity judgments between the images. However, which exact target knowledge was activated and consequently prompted assimilation and contrast remains in the realm of the hypothetical.

In our studies, we showed the target and context image next to each other in a simultaneous presentation. According to Wedell et al. (1987), the simultaneous evaluation of two stimuli makes it difficult for the perceiver to distinguish the subjective experiences evoked by each stimulus. The evaluation of the target image should therefore shift toward the evaluation of the context stimulus seen in juxtaposition compared to when the target is evaluated after the context stimulus. This would result in an assimilation effect. Our data from the negative, valence first condition in Study 1 as well as the negative conditions in Study 2 are in accordance with Wedell et al. (1987). However, in the other conditions, we did not find any assimilation, even though the two images were presented simultaneously. We argue that simultaneous presentation may facilitate an assimilation effect, but this does not lead per se to an approximation of the stimuli.

Limitations and Future Work

In the following, we review the main limitations of our study and discuss possibilities for future work. In our studies, we showed that contrast and assimilation effects were influenced by the order of the evaluative dimensions, but we could not say with certainty which target knowledge was activated. Future studies should investigate the role of evaluation dimensions and their influence on the emotional interpretation of art. In our experiment, rating ambiguity in general had an effect on the valence ratings in our first study. But the intensity of perceived ambiguity in the target images themselves did not show a clear correlation with the valence ratings in the single and the paired conditions. This could be due to the fact that target images used in our experiment were rated as moderately ambiguous. In future work, using images that exhibit more pronounced ambiguity could reveal an answer to the question of whether images depicting a rather vague meaning are more susceptible to contextual influences (Herr et al., 1983; Muth et al., 2015). We also investigated whether formal similarity (Arnheim, 1957) would enhance the influence of the context image on the target image, but we did not find any evidence for this. On the contrary, the effect was more pronounced in the condition with no formal similarity. This result raises the question of whether the content of images has a greater influence on neighboring images than their formal aspects. However, since the images used in our studies were not found to be extremely similar, this question requires further experiments with images whose formal similarity is more pronounced.

We also investigated viewer-centered aspects. Since trained designers are accustomed to evaluating the formal aspects of images, we expected the influence of formal similarity to be more pronounced in experts' ratings than in laypeople's ratings (Rota & Zellner, 2007). Although our data showed that there was no difference between experts and laypeople in terms of formal similarity and in subsequent valence ratings, these results need to be examined in more detail using comparable group sizes. This lack of clear results may also be due to the fact that the images were not rated to be as formally similar as we expected them to be. We also expect that measuring the current affective state of the viewers not only in the middle of the experiment but also at the end of the experiment could have shed more light on the question why we found a more pronounced negative effect in the rating of the juxtaposed images in our studies (Konečni & Sargent-Pollock, 1977; Leder et al., 2004). Although the participants' current affects showed no significant negative moods, it is possible that it changed during the course of the study or even due to the evaluation of the images. Future studies should therefore assess participants' affect also at the end of the study, especially if the rating dimensions concern valence.

Conclusion

What insights do these results offer to visual designers and curators? In summary, our results empirically support experts' tacit knowledge in placing images. The image context influences the emotional interpretation of an image. This confirms the importance of the work of curators and designers in visual communication in arranging images, because context influences the affective interpretation we attribute to images. A curator can render the affective interpretation of an image more negative by placing it next to a similar image with a negative connotation. A designer makes an image look more positive by displaying it in the context to a negative image that shows no similarities to it. Essentially, our results confirm that it is highly desirable to show artworks again and again in different contexts in order to enable new experiences that allow us to rediscover different, contradictory facets of an image. Interestingly, our results also show that other, nonpictorial factors influence our evaluations of images, which is consistent with a growing body of work demonstrating the variety of contextual influences on art perception (Pelowski & Specker, 2020). If a cognitive context in which an art image is viewed is made salient, it will have an influence on the evaluation of the image. In that sense, curators and designers can effectively use not only neighboring images as an influencing context but also the thematic orientation of an exhibition or the title of a catalogue to make (art) images appear in a new, undiscovered light.

References

Allesch, C. G. (2018). *Fechner: Vorschule der Ästhetik*. Springer Spektrum. https://doi.org/10.1007/978-3-662-57313-6

Arielli, E. (2012). Contrast and assimilation in aesthetic judgments of visual artworks. *Empirical*

Studies of the Arts, *30*(1), 59–74. https://doi.org/10.2190/EM.30.1.f

Arnheim, R. (1957). *Art and visual perception: A psychology of the creative eye*. University of California Press.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–732.

Bååth, R. (2015). *Bayesian First Aid*. https://github.com/rasmusab/bayesian_first_aid

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, *20*(5), 286–290. https://doi.org/10.1177/0963721411422522

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Bradley, M. M., Keil, A., & Lang, P. J. (2012). Orienting and emotional perception: Facilitation, attenuation, and interference. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00493

Bradley, M., Sabatinelli, D., Lang, P., Fitzsimmons, J., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral neuroscience*, *117*, 369–380. https://doi.org/10.1037/0735-7044.117.2.369

Calbi, M., Heimann, K., Barratt, D., Siri, F., Umiltà, M. A., & Gallese, V. (2017). How context influences our perception of emotional faces: A behavioral study on the Kuleshov effect. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01684

Chang, S., Kim, C.-Y., & Cho, Y. S. (2017). Sequential effects in preference decision: Prior preference assimilates current preference. *PLOS ONE*, *12*(8). https://doi.org/10.1371/journal.pone.0182442

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cohn, N. (2013). Visual narrative structure. *Cognitive Science*, *37*(3), 413–452. https://doi.org/10.1111/cogs.12016

Cohn, N. (2015). Climbing trees and seeing stars: Combinatorial structure in comics and diverse domains. In I. Toivonen, P. Csúri, & E. Van Der Zee (Eds.), *Structures in the mind* (pp. 379–392). MIT Press.

Comtois, D. (2021). *summarytools: Tools to quickly and neatly summarize data*. https://CRAN.R-project.org/package=summarytools

Cupchik, G. C., Shereck, L., & Spiegel, S. (1994). The effects of textual information on artistic communication. *Visual Arts Research*, *20*, 62–78.

Cupchik, G., Winston, A. S., & Herz, R. (1992). Judgments of similarity and difference between paintings. *Visual Arts Research*, *18*, 37–50.

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference*. Macmillan International Higher Education.

Dolese, M., Zellner, D. A., Vasserman, M., & Parker, S. (2005). Categorization affects hedonic contrast in the visual arts. *Bulletin of Psychology and the Arts*, *5*(2), 21–25.

Flacke, M. (2016). Ausstellen als Narration. In M. Walz (Ed.), *Handbuch Museum: Geschichte, Aufgaben, Perspektiven* (pp. 253–257). J. B. Metzler. https://doi.org/10.1007/978-3-476-05184-4_56

Förster, J., Liberman, N., & Kuschel, S. (2008). The effect of global versus local processing styles on assimilation versus contrast in social judgment. *Journal of Personality and Social Psychology*, *94*(4), 579–599. https://doi.org/10.1037/0022-3514.94.4.579

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics*, *40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, *5*(supplement), 123–147. https://doi.org/10.1177/10298649020050S105

Ganz, D., & Thürlemann, F. (2010). Zur Einführung. In D. Ganz & F. Thürlemann (Eds.), *Das Bild im Plural* (pp. 7–40). Dietrich Reimer Verlag.

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. https://doi.org/10.1016/j.actpsy.2015.01.005

Gaver, W. W., Beaver, J., & Benford, S. (2003). Ambiguity as a resource for design. *NEW HORIZONS*, *5*, 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. https://doi.org/10.1016/j.actpsy.2014.06.008

Gibson, J. J. (1978). The ecological approach to the visual perception of pictures. *Leonardo*, *11*(3), 227. https://doi.org/10.2307/1574154

Gibson, R., & Zillmann, D. (2000). Reading between the photographs: The influence of incidental pictorial information on issue perception. *Journalism & Mass Communication Quarterly*, *77*(2), 355–366. https://doi.org/10.1177/107769900007700209

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*(4), 323–340.

Higgins, E. T., & Chaires, W. M. (1980). Accessibility of interrelational constructs: Implications for stimulus encoding and creativity. *Journal of Experimental Social Psychology*, *16*(4), 348–361. https://doi.org/10.1016/0022-1031(80)90027-X

Hofmann, A. (1985). Thoughts on the study and making of visual signs, Basle School of Design/Yale School of Art, 1947 to 1985. *Design Quarterly*, *130*, 1–20.

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.

https://doi.org/10.1080/17470210903038974

Jakesch, M., Leder, H., & Forster, M. (2013). Image ambiguity and fluency. *PLOS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0074084

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191–213. https://doi.org/10.1177/102986490601000203

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

Kondo, A., Takahashi, K., & Watanabe, K. (2012). Sequential effects in face-attractiveness judgment. *Perception*, *41*(1), 43–49.

Konečni, V. J., & Sargent-Pollock, D. (1977). Arousal, positive and negative affect, and preference for Renaissance and 20th-century paintings. *Motivation and Emotion*, *1*(1), 75–93. https://doi.org/10.1007/BF00997582

Kruschke, J. K., & Meredith, M. (2021). *BEST: Bayesian estimation supersedes the t-test*. https://CRAN.R-project.org/package=BEST

Kuleshov, L. V. (1974). *Kuleshov on film: Writings*. University of California Press.

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

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*, *1*, 39–58.

Lang, P. J., Simons, R. F., Balaban, M., & Simons, R. (2013). *Attention and orienting: Sensory and motivational processes*. Psychology Press.

Lauring, J. O., Pelowski, M., Forster, M., Gondan, M., Ptito, M., & Kupers, R. (2016). Well, if they like it... Effects of social groups' ratings and price information on the appreciation of art.

Psychology of Aesthetics, Creativity, and the Arts, *10*(3), 344–359. https://doi.org/10.1037/aca0000063

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., 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. https://doi.org/10.1080/02699931.2013.870132

Marian, D. E., & Shimamura, A. P. (2012). Emotions in context: Pictorial influences on affective attributions. *Emotion*, *12*(2), 371–375. https://doi.org/10.1037/a0025517

McLain, D. L. (2009). Evidence of the properties of an ambiguity tolerance measure: The Multiple Stimulus Types Ambiguity Tolerance scale–II (MSTAT–II). *Psychological Reports*, *105*(3), 975–988. https://doi.org/10.2466/PR0.105.3.975-988

Mobbs, D., Weiskopf, N., Lau, H. C., Featherstone, E., Dolan, R. J., & Frith, C. D. (2006). The Kuleshov effect: The influence of contextual framing on emotional attributions. *Social Cognitive and Affective Neuroscience*, *1*(2), 95–106. https://doi.org/10.1093/scan/nsl014

Mullennix, J. W., Kristo, G. M., & Robinet, J. (2018). Effects of preceding context on aesthetic preference. *Empirical Studies of the Arts*, *38*(2), 149–171. https://doi.org/10.1177/0276237418805687

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*(3), 472–489. https://doi.org/10.1037/0033-295X.110.3.472

Mussweiler, T., & Bodenhausen, G. V. (2002). I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology*, *82*(1), 19–32. https://doi.org/10.1037/0022-3514.82.1.19

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. https://doi.org/10.1037/a0038814

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, *72*(6), 407–418. https://doi.org/10.1037/h0022602

Pedersen, T. L. (2020). *patchwork: The composer of plots*. https://CRAN.R-project.org/package=patchwork

Pegors, T. K., Mattar, M. G., Bryan, P. B., & Epstein, R. A. (2015). Simultaneous perceptual and response biases on sequential face attractiveness judgments. *Journal of Experimental Psychology: General*, *144*(3), 664–673. https://doi.org/10.1037/xge0000069

Pelowski, M., Gerger, G., Chetouani, Y., Markey, P. S., & Leder, H. (2017a). But is it really art? The classification of images as “art”/“not art” and correlation with appraisal and viewer interpersonal differences. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01729

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. https://doi.org/10.1016/j.plrev.2017.02.003

Pelowski, M., & Specker, E. (2020). The general impact of context on aesthetic experience. In M. Nadal & O. Vartanian (Eds.), *Empirical aesthetics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.42

Pelowski, M., Specker, E., Gerger, G., Leder, H., & Weingarden, L. S. (2020). Do you feel like I do? A study of spontaneous and deliberate emotion sharing and understanding between artists and perceivers of installation art. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(3), 276–293. https://doi.org/10.1037/aca0000201

Price, V., Tewksbury, D., & Powers, E. (1997). Switching trains of thought: The impact of news frames on readers' cognitive responses. *Communication Research*, *24*(5), 481–506.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reymond, C. (2013). Bild und Bild: Eine Untersuchung zwischen-bildlicher Prozesse. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, *23*, 1–23. https://doi.org/10.21240/mpaed/23/2013.06.04.X

Reymond, C., Pelowski, M., Opwis, K., Takala, T., & Mekler, E. D. (2020). Aesthetic evaluation of digitally reproduced art images. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.615575

Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social Cognitive and Affective Neuroscience*, *3*(3), 270–278. https://doi.org/10.1093/scan/nsn021

Rodriguez, L., & Dimitrova, D. V. (2011). The levels of visual framing. *Journal of Visual Literacy*, *30*(1), 48–65. https://doi.org/10.1080/23796529.2011.11674684

Rota, L. M., & Zellner, D. A. (2007). The categorization effect in hedonic contrast: Experts differ from novices. *Psychonomic Bulletin & Review*, *14*(1), 179–183. https://doi.org/10.3758/BF03194047

Rozin, P., & Royzman, E. B. (2001). Negativity bias, Negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Samara, T. (2007). *Design elements: A graphic style manual*. Rockport Publishers.

Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593–611.

https://doi.org/10.1080/02699930341000239

Schwarz, N., & Bless, H. (1991). *Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment* (ZUMA-Arbeitsbericht 1991/05). Zentrum für Umfragen, Methoden und Analysen -ZUMA-.

Szubielska, M., & Sztorc, A. (2019). The influence of extended contextual information provided in a contemporary art gallery on aesthetic experience of art faculties students. *Polish Psychological Bulletin*, *50*(4), 345–351.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

Tousignant, C., & Bodner, G. E. (2014). Context effects on beauty ratings of photos: Building contrast effects that erode but cannot be knocked down. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 81–86. https://doi.org/10.1037/a0034942

Tousignant, C., & Bodner, G. E. (2018). Context effects on beauty ratings of abstract paintings: Contrast, contrast, everywhere! *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 369–377. https://doi.org/10.1037/aca0000146

Unipark. (2017). *Questback GmbH*. EFS Survey. https://www.unipark.com/

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. https://doi.org/10.1037/0033-2909.134.3.383

Wedell, D. H., Parducci, A., & Geiselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology*, *23*(3), 230–249. https://doi.org/10.1016/0022-1031(87)90034-5

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Zellner, D. A., Mattingly, M. C., & Parker, S. (2009). Categorization reduces the effect of context on hedonic preference. *Attention, Perception, & Psychophysics*, *71*(6), 1228–1232. https://doi.org/10.3758/APP.71.6.1228

Zellner, D. A., Rohm, E. A., Bassetti, T. L., & Parker, S. (2003). Compared to what? Effects of categorization on hedonic contrast. *Psychonomic Bulletin & Review*, *10*(2), 468–473. https://doi.org/10.3758/BF03196508

Naming Images in Aphasia:

Effects of Graphic Representations and Photographs on Naming Performance in People With and Without Aphasia

Claire Reymond[1]*, Sandra Widmer Beierlein[2], Christine Müller[1], Ricarda Reutimann[1], Katrin P. Kuntner[2], Noelia Falcon Garcia[2], Indre Grumbinaite[1], Simone Hemm[3], Markus Degen[3], Fabrizio Parillo[3], Stefan Karlin[3], Sunghea Park[2], Michael Renner[1], and Anja Blechschmidt[3]

FHNW University of Applied Sciences and Arts Northwestern Switzerland,

[1]Academy of Art and Design FHNW, [2]School of Education FHNW, and [3]School of Life Sciences FHNW.

Author Note

Claire Reymond https://orcid.org/0000-0002-8988-7263

Sandra Widmer Beierlein https://orcid.org/0000-0002-8054-2365

Christine Müller https://orcid.org/0000-0001-8235-060X

Ricarda Reutimann https://orcid.org/0000-0003-2236-3809

Katrin P. Kuntner https://orcid.org/0000-0003-4085-5428

Noelia Falcon Garcia https://orcid.org/0000-0002-4490-4250

Indre Grumbinaite https://orcid.org/0000-0001-6211-8459

Simone Hemm https://orcid.org/0000-0002-5726-6761

Markus Degen https://orcid.org/0000-0001-7009-782X

Fabrizio Parrillo https://orcid.org/0000-0002-9998-1317

Stefan Karlin https://orcid.org/0000-0002-2614-9407

Sunghae Park https://orcid.org/0000-0003-1163-3199

Michael Renner https://orcid.org/0000-0001-9607-372X

Anja Blechschmidt https://orcid.org/0000-0003-3312-7529

The authors have no known conflict of interest to disclose and thank the FHNW for funding E-Inclusion, the study participants for their motivated participation, and the speech therapy teams and departments of the practice partners for their dedicated collaboration: REHAB Basel, Reha Rheinfelden, Bad Schinznach (Privat-Klinik Im Park); Cantonal Hospitals Aarau, Baden, Baselland, Graubünden and Olten (Solothurn Hospitals); Hospitals Schaffhausen and Zofingen; University Hospital Zurich (Clinic for Neurology) and University Geriatric Medicine FELIX PLATTER; Stadtspital Waid and Triemli; speech therapy practices Fanny Dittmann-Aubert, Isabelle Facchini-Baumann, LogoTreffpunkt, mund'art, rundum therapie, Petra Dietiker and Unterstrass.

The authors would also like to thank the team of designers from the Academy of Art and Design FHNW: Lena Meier, Sarah Schiltknecht, Laurent Poffet, Alisa Strub, and Karen Trachsel.

*Correspondence concerning this article should be addressed to Claire Reymond, Institute Digital Communication Environments, Academy of Art and Design, University of Applied Sciences and Arts Northwestern Switzerland, Freilager-Platz 1, CH-4002 Basel, Switzerland. E-mail: claire.reymond@fhnw.ch

Naming Images in Aphasia:
Effects of Graphic Representations and Photographs on Naming Performance in People With and Without Aphasia

Background: Picture naming is a common tool in aphasia diagnosis and therapy. However, opinions differ as to which type of image (e.g., photographs, drawings) is most suitable for naming tasks and whether there is a difference on naming accuracy and latencies based on image type. Moreover, recent studies have mainly analyzed color photographs and black-and-white line drawings leaving out image types like graphic representations that apply beneficial image features like color, controlled size, or texture.
Aims: To shed more light on appropriate image types for people with aphasia, we created graphic representations depicting nouns and verbs and compared them to photographic stimuli in a naming task including persons with aphasia (PWA) and a control group (CG).
Methods & Procedures: 33 PWA and 33 age matched CG participated in the study. Naming correctness and latencies were measured in two different conditions: concepts depicted as colored photographs vs. as graphic representations. 128 pictures of linguistically controlled German-language concepts (64 nouns, 64 verbs) had to be named. The designed graphic stimuli were developed by professional designer based on photographs. The photographs were selected from stock image databases according to a defined image concept. This image concept was based on empirical findings regarding the positive effect of favorable image features (e.g., color, texture, shading) and was applied to the selection of the photographs as well as to the creation of the graphic representations. The images were presented in pseudo-randomized sequences on a tablet and all reactions were videotaped. The data from the main study was analyzed using generalized linear mixed models (GLMM).
Outcomes & Results: Our analysis showed no significant difference in naming correctness and latencies between photographic and graphic stimuli. Comparing graphic representations and photographs in PWA when naming nouns and verbs, graphic stimuli were named more correctly on average, however, this difference was below significance level.
Conclusions: In our study we showed that graphic representations, when including favorable image features like color, texture and shading, can evoke the same naming performance as photographs. We therefore advocate the use of graphic representations that include favorable image features. These can be used in combination with photographs in an image set, especially when depicting concepts that benefit from reduced representation, e.g., verbs.

Keywords: aphasia; picture naming; image type; image features; E-Inclusion

Introduction

Naming images is a common paradigm for investigating unimpaired (Bates et al., 2003; Glaser, 1992; Kohnert, 2004) and impaired language processing (e.g., Cotelli et al., 2007; Cuetos et al., 2005; Howard et al., 1985; Kohn & Goodglass, 1985; Laine et al., 1997). Naming objects or scenes depicted on pictures is also used to diagnose the severeness of aphasia using standardized and psychometrically validated language tests (e.g., Aachener Aphasie Test; Huber et al., 1983; Boston Naming Test; Kaplan et al., 2001; Bielefelder Aphasie Screening; Richter et al., 2006; Quick Aphasia Battery; Wilson et al., 2018) or to train and re-learn language in aphasia therapy (e.g., Brumbi et al., 2017; Pfab et al., 2015; Stark, 1992). Despite the longstanding use of images in aphasia diagnostics and therapy, there is still disagreement on what relevant images for persons with aphasia actually should look like (Brown & Thiessen, 2018). In this paper, the term *image type* refers to the technique used to create an image and categorizes an image as, for example, drawing, photograph, or painting. Each image type can be further differentiated: For example, drawings can vary in terms of their level of generalization (from as detailed as possible like in scientific illustrations to simplified as in pictograms). Different image types in the context of aphasia—usually classified as *drawn* or *photographic* stimuli—have been investigated regarding their effect on naming performance. *Drawings* mostly refer to a heterogeneous group of black-and-white or colored line drawings. In line drawings the depicted figure is outlined with a black line presented on a white background, in colored line drawings the surfaces are colored. *Pictograms*, *pictographs* or *icons* are often described as a sub-category of drawn stimuli and are mostly used interchangeably for the same image type: a black graphic form that conveys meaning through its pictorial resemblance to a physical object. *Photographs* are colored or black-and-white photographic images, showing objects or persons embedded in a scene or cropped and shown on a plain background.

Comparing object recognition in color photographs and black-and-white line drawings in healthy participants, Heuer (2016) showed that object recognition was facilitated by color photographs. Based on this, she argued that the image types used in aphasia diagnostics should be examined in more detail. But while some older studies have demonstrated differences of image type on naming performance (Benton et al., 1972; Bisiach, 1966), others have shown no such effects (Corlew & Nation, 1975). The current heterogeneity of available picture stimuli and their effects on naming performance places a unique challenge for practitioners in aphasia therapy when it comes to providing appropriate image material to their patients. The vast variety of pictures to choose from might also cause uncertainty as to which types of images should be used in aphasia diagnostics (Brown & Thiessen, 2018). The aim of our paper is therefore to clarify the role of image type on naming performance in persons with aphasia (PWA) and adults without aphasia (control group: CG). We do so by comparing naming responses (naming correctness and response latencies) to two image types, which we created or selected from professional image databases.

Effects of image type on naming performance and communication behavior in PWA

Drawn stimuli may be regarded as a translation of objects, persons or scenes seen in the real world. As such, drawings offer the possibility to focus on core features of an object or scene, accentuating the meaning of the image and leaving out unnecessary details. Specifically line drawings are said to facilitate text comprehension in normal (Readence & Moore, 1981) and impaired readers (Doak et al., 1996; Mayer & Villaire, 2007) because keeping the picture simple in construction seems to be better suited to enhance comprehension than artistic or overly detailed images (Readence & Moore, 1981, p. 222). Line drawings are also considered beneficial for evoking language in PWA (Kagan & LeBlanc, 2002; Pound et al., 1999) and are therefore traditionally used in diagnostic tests (e.g., Huber et al., 1983; Richter et al., 2006). When studying objects depicted as colored line drawings, black-and-white line drawings, and black-and-white line drawings with lines drawn across the picture, Bisiach (1966) compared recognition and naming accuracy in nine PWA. There was no significant difference in recognition of the images. However, the ratio of correct naming based on the correctly recognized images was significantly higher for the colored line drawings compared to the black-and-white and the crossed-out images. Benton et al. (1972) compared small and large black-and-white line drawings of objects and real objects on naming accuracy in 18 PWA. A significant difference in naming accuracy between line drawings and objects was observed—although the naming performance in the different conditions were extremely close to each other. The authors suggest that real objects may provide more opportunity for association and semantic activation than line drawings which may in turn favor retrieval of the object' name. The size of line drawings did not influence naming accuracy. Bisiach (1966) and Benton et al. (1972) showed that proximity to the real object (as a detailed color drawing or as a real object) has a positive influence on naming performance. In contrast to those results, Corlew and Nation (1975) showed no differences based on different stimulus types. They investigated the correct naming of real objects compared to reduced black and white line drawings of the same objects. The naming responses of 14 PWA did not differ significantly based on the stimuli.

Photographs depict objects, scenes or persons similarly to how they are perceived in everyday life. In this sense, photographs show what is depicted with all its rich details which facilitates object recognition and may therefore be beneficial for language processing in PWA (Heuer, 2016). When comparing the effect of colored and black-and-white photographs of objects on naming latency and correctness, Mohr (2010) showed that 29 PWA as well as 60 adults without aphasia named the objects on the color photographs significantly faster and more accurately than the black-and-white photographs.

The effectiveness of photographs to support communication in PWA has been shown in studies comparing photographs with line drawings and/or pictograms. Analyzing communication behavior with the support of photographed objects and colored pictograms of objects, Ho et al. (2005) described that their two participants with global aphasia communicated about more topics and had fewer communication interruptions when supported by both image types than when no images were present. Additionally, more pointing behavior was observed with photographs than with pictograms. When testing the use of augmented and alternative communication devices, Griffith et al. (2014) showed that four PWA more frequently used personally relevant photographs than colored line drawings when retelling a narrative. However, participants orally reported both image types to be equally helpful. Ma et al. (2009) investigated different types of images (colored and black-and-white photographs, black-and-white pictograms) to icons in 50 PWA on their effectiveness in communication replacement. In this study, icons denote pixelated black-and-white line drawings with and without color. Icons and images were shown to be equally functional in conveying information and thus equally suitable for communicating single words (Ma et al., 2009).

Effects of image features on naming performance for objects

To name an object on an image correctly it needs to be recognized correctly first (Alario et al., 2004; Heuer, 2016; Humphreys et al., 1999). At the same time a variety of image features influence recognition and naming performance in persons without neurological deficits (for an overview see Johnson et al., 1996). Image features describe formal characteristics of an image such as the color used or the contrast within an image. The size of the depicted object affects the time needed to name an object: Depicted objects that are smaller than 4° to 6° of visual angle may be difficult to perceive because they are too small. Very large images, in turn, require eye movement for optimal foveal perception, increasing the difficulty of identification (Biederman & Cooper, 1992; Johnson et al., 1996, p. 117). Maintaining a realistic size ratio within an object category was considered in the image set of Snodgrass and Vanderwart (1980). The objects' color is relevant when it underlines the prototypicality of the object. When color-prototypical objects are presented in incongruent colors, they are named slower than when shown in their prototypical color (Naor-Raz et al., 2003; Rossion & Pourtois, 2001; Therriault et al., 2009; Wurm et al., 1993). A medium color contrast of the object has been shown to provoke faster naming latencies than low color contrast (Brodie et al., 1991, Experiment 1). The depicted texture of an object (depiction of the materiality of the object) and shading (dark and light colored surfaces on an object simulating three-dimensionality) facilitates naming performance (Adlington, 2009; Brodie et al., 1991; Rossion & Pourtois, 2001). The view on objects also plays a role in naming performance: When objects are displayed in the canonical view, that is, in the position where the information identifying the object is optimally visible (e.g., seeing a zebra from the side favors recognition compared to when it is seen from above), naming of the object is fastest (Palmer, 1981; Snodgrass & Vanderwart, 1980).

Depictions of verbs

It is known that naming verbs is more demanding than naming objects, for persons without neurological deficits as well as for PWA (Bastiaanse et al., 2003; Bastiaanse &

Jonkers, 1998; De Bleser & Kauschke, 2003; Kim & Thompson, 2000). This is based on factors like differences in semantic representation (Vinson & Vigliocco, 2002) or the higher morphological complexity of verbs (Vigliocco et al., 2006). In addition, the imageability of verbs poses a challenge compared to the imageability of nouns (Bird et al., 2003) whereas “imageability” is defined as the “extent to which an object name evokes few or many different images for a particular object” (Alario et al., 2004, p. 141). Bird et al., (2003) investigated the effect of imageability in four PWA and showed that if the depicted verbs were controlled for imageability no differences in reading or writing were found for verbs and nouns. This suggests that the PWA found verbs to be more difficult because of the type of representation, rather than the word class. Where depicting concrete nouns allows for a clear boundary of the figure on a plain background, verbs are characterized by the need to show actors in a more or less complex context. This fundamental difficulty is reflected in the comparably small number of image sets that depict verbs (for image sets containing depictions of verbs, see Fiez & Tranel, 1997; Khwaileh et al., 2018; Masterson & Druks, 1998; Székely et al., 2004). The set of 280 images by Fiez and Tranel (1997) shows verbs as grayscale photographs of “persons, animals, and objects engaged in ongoing actions” (p. 547) on a plain or structured background. Khwaileh et al. (2018), Masterson and Druks (1998) and Székely et al. (2004) provided sets of depicted verbs by assembling black-and-white line drawings of persons, animals or objects performing different actions. Akinina et al. (2015) have provided a set of 375 black-and-white as well as a colored version of line drawings depicting action pictures and verbs. Investigating the depiction of verbs, Thiessen et al. (2016) have shown that task-engaged depiction (the depicted person turns towards the performed activity) draws significantly more visual attention of PWA as well as adults without aphasia to the objects than when the depicted person looks into the camera. This in turn might have a positive influence on the naming of activities.

In summary, the few studies exerted on the effect of different image types on naming performance when naming objects have produced mixed results. Some studies have indicated that if photographs were compared to reduced black-and-white line drawings, they evoked better naming performance. However, there is evidence that specific image features like controlled size, color, texture and viewpoint on the object affect naming performance of objects in persons without neurological deficits (Biederman & Cooper, 1992; Johnson et al., 1996; Naor-Raz et al., 2003; Palmer, 1981; Therriault et al., 2009). Thus, comparing color photographs with black-and-white drawings might restrict the positive effect of these features—all of which can be applied to drawings—on naming performance in PWA. To our knowledge, those image features have hardly been systematically implemented into image material depicting objects and verbs specially produced for PWA. Moreover, the comparably small number of image sets providing depictions of verbs have mainly used grayscale photographs, black-and-white or colored line drawings. The predominant use of line drawings for depicting verbs in the existing image sets might suggest that a schematic and reduced image type may be beneficial for representing verbs.

Studies that have actually investigated the role of image types on naming performance in PWA and not merely in persons without aphasia is still quite small. To our knowledge, the effect of drawn stimuli that go beyond reduced (black-and-white) line drawings and implement favorable image features like color, texture, and shading on naming correctness and latencies in PWA has hardly been researched so far and is the aim of our study[1].

Research Questions of the Present Study

In our study, we developed drawn stimuli incorporating image features that have been shown to have a favorable effect on naming performance, e.g. controlled size, color, perspective, and texture (Adlington, 2009; Biederman & Cooper, 1992; Palmer, 1981; Rossion & Pourtois, 2001; Wurm et al., 1993). The type of stimuli we compared to color photographs in our study were drawn on the computer using Adobe Illustrator and Adobe Photoshop (https://www.adobe.com). As such we consider our digital drawings to be a sub-category of the image type of drawings and refer to them as *graphic representations*. We compared naming correctness and response latencies between these graphic representations and photographs depicting nouns and verbs in a picture naming task in a group of 33 PWA and an age-matched control group of 33 participants.

This study was designed to assess whether photographic and graphic images affect naming correctness and naming latency of nouns and verbs for PWA. We also compared naming performance in PWA and the CG.

Materials and Methods

The study was approved by the Ethics Committees of Northwestern and Central Switzerland (EKNZ) and the Cantonal Ethics Committee Zurich (kek.zh.ch), Project ID: 2019-00084. The study was also registered at ClinicalTrials.gov, Identifier: NCT05164380.

[1] For in depth information on the interdisciplinary project *E-Inclusion* in which this study was embedded see Widmer Beierlein et al. (2021).

Participants

A total of 33 adults with aphasia (M_{age} = 58.1, SD = 13.6; 18 males, 15 females) were included in our study. The approximate number of participants was previously defined by a power analysis based on simulation. PWA were recruited in cooperation with local speech therapists, hospitals and rehabilitation clinics. PWA had to be diagnosed with minimal, light or moderate aphasia by the Aachener Aphasie Test (AAT; Huber et al., 1983). All of them were in the post-acute or chronic phase with a minimum of six weeks post onset. Twenty PWA were diagnosed with anomic aphasia, five PWA with Wernicke aphasia, seven with Broca aphasia and one participant with global aphasia. In 26 persons the aphasia was caused by an ischemic and in three by a hemorrhagic stroke, two had a tumor and two a craniocerebral trauma. All participants had to show sufficient language comprehension to follow the instructions during the experiment, an attention span of minimum 45 minutes, a mild dysarthria or apraxia of speech was allowed but no indication for dementia according to the treating speech therapist. Further they had to have intact color vision, normal or corrected to normal vision and hearing attested on a questionnaire filled out by the patient before study enrollment. Refer to Table 1 for description of PWA.

The age-matched control group consisted of 33 participants (M_{age} = 58.2, SD = 14.2; 16 males, 17 females) all with intact color vision, normal or corrected to normal vision and hearing and no neurological diseases in their health history. They were recruited via inquiries sent to retiree associations, the researchers' circle of acquaintances as well as by asking relatives of participants with aphasia. A two-sided t-test for related samples (due to matching) showed no significant difference regarding age between individuals with and without aphasia ($T(32) = 0.020$, $p = .984$).

Table 1. Description of PWA.

ID	Age	Gender	Handedness	Education [years]	First language	Aphasia type	Aphasia phase	Time since onset [months]	AAT Severity	AAT Severity naming	Aetiology	Lateralisation of damage
1	74	F	Right	10	Swiss German	Anomic	Post-acute	9	Light	Light	Ischemic stroke	Left
2	48	F	Right	12	Swiss German	Anomic	Chronic	40	Light	Light	Ischemic stroke	Left
3	53	F	Right	12	Swiss German	Anomic	Post-acute	10	Minimal	Light	Braintumor	Left
4	60	M	Right	13	Swiss German	Anomic	Post-acute	6	Light	Light	Ischemic stroke	Left & right
5	72	M	Right	16	Swiss German	Anomic	Post-acute	7	Light	Minimal	Ischemic stroke	Left
6	62	F	Left (right for writing)	12	Swiss German	Wernicke	Chronic	16	Minimal	Medium	Ischemic stroke	Left
7	43	M	Right	13	Swiss German	Anomic	Chronic	16	Light	Light	Ischemic stroke	Left
8	36	F	Left (previously right)	12	Swiss German	Anomic	Chronic	108	Minimal	Light	Ischemic stroke	Left
9	30	F	Left	9	Swiss German	Broca	Chronic	130	Medium	Light	Ischemic stroke, Craniocerebral trauma	Left
10	47	M	Right	16	Swiss German	Anomic	Chronic	17	Minimal	Minimal	Craniocerebral trauma	Left
11	55	M	Right	13	Swiss German	Global	Chronic	217	Severe	Medium	Ischemic stroke	Left
12	74	M	Two-handed	12	Swiss German	Broca	Chronic	61	Medium	Medium	Hemorrhagic stroke	Left
13	27	M	Right	14	Swiss German	Anomic	Chronic	105	Light	Light	Ischemic stroke	Left
14	69	F	Right	10	Swiss German	Anomic	Post-acute	1	Light	Minimal	Braintumor	Left
15	43	M	Left	29	Swiss German	Wernicke	Chronic	43	Severe	Light	Ischemic stroke	Left
16	67	M	Right	13	Swiss German	Anomic	Chronic	28	Minimal	Minimal	Ischemic stroke	Right
17	75	F	Right	9	Swiss German	Anomic	Chronic	25	Minimal	Minimal	Ischemic stroke	Left
18	60	M	Right	9	Swiss German	Anomic	Chronic	15	Minimal	Minimal	Ischemic stroke	Left
19	69	M	Right	19	Swiss German	Anomic	Post-acute	2	Light	Light	Ischemic stroke	Left
20	58	F	Right	12	Swiss German	Anomic	Post-acute	5	Minimal	Light	Hemorrhagic stroke	Right
21	42	F	Left	11	Swiss German	Broca	Chronic	140	Medium	Medium	Hemorrhagic stroke	Left
22	54	M	Two-handed	15	Swiss German	Broca	Chronic	120	Light	Medium	Ischemic stroke	Left
23	58	F	Right	13	Swiss German	Broca	Chronic	34	Medium	Medium	Ischemic stroke	Left
24	56	M	Right	16	Swiss German	Anomic	Chronic	65	Minimal	Light	Ischemic stroke	Left
25	84	F	Right (retrained to left)	10	Swiss German	Anomic	Chronic	258	Light	Light	Ischemic stroke	NA*
26	53	F	Right	1	Swiss German	Anomic	Chronic	142	Light	Light	Hemorrhagic stroke	Left
27	48	F	Right	17	Swiss German	Wernicke	Chronic	138	Medium	Medium	Ischemic stroke	Left
28	65	M	Right	16	Swiss German	Broca	Chronic	91	Minimal	Light	Ischemic stroke	Left
29	75	F	Right	10	Swiss German	Anomic	Chronic	78	Minimal	Light	Ischemic stroke	NA*
30	68	M	Right	9	Swiss German	Wernicke	Chronic	16	Light	Medium	Ischemic stroke	Left
31	62	M	Right	12	Swiss German	Wernicke	Post-acute	2	Medium	Medium	Ischemic stroke	Left
32	63	M	Right	9	Swiss German	Anomic	Chronic	14	Light	Light	Ischemic stroke	Left
33	67	M	Left	15	Swiss German	Broca	Chronic	124	Medium	Medium	Ischemic stroke	Left

- This information is missing because these patients did not provide any information.

Stimuli

We created graphic representations of 128 words—64 verbs and 64 nouns to be compared to 128 photographs in our study. Eight trial items (four nouns, four verbs) were added to be used as examples before the main test items. The set of graphic representations can be seen and downloaded on the following platform: https://mediathek.hgk.fhnw.ch/einclusion/. The images may be used by citing the present article as a reference (See Fig. 1 as an example for the style of graphic representation).

Figure 1: Graphic representation of a drill, showing the use of shading to provoke the illusion of three-dimensionality and different color tones to show the texture of the object.

Selection of words

Naming correctness and naming latencies are influenced by several linguistic aspects such as word frequency, word length, accent, image agreement and name agreement (Adlington, 2009; Alario et al., 2004; Barry et al., 1997; Bates et al., 2003; Kemmerer, 2014; Laiacona et al., 2001; Levelt et al., 1999; Menn & Bastiaanse, 2016; Meyer et al., 2003; Snodgrass & Vanderwart, 1980). To assemble a comparable image set, each German-language term needed to fulfill linguistic specifications: all of them had two syllables, a trocheic rhythm with accent on the first syllable, maximally two morphemes, and were low frequent according to “Swiss German vocabulary” (Leipzig Corpora Collection). Only active and full verbs and nouns displaying object names (no proper nouns or colors) were included and the concepts had to be imageable.

Graphic Representations & Photographs

The graphic representations were developed within the research team and created by trained designers. They were designed based on photographs chosen according to the defined image concept and bought on stock images platforms (iStock www.istockphoto.com and Getty Images www.gettyimages.ch) (Reymond et al., 2019). The following image features were defined as image concept and applied identically for photographic stimuli and graphic representations: Color was used to emphasize the characteristics of the subject depicted and natural objects were shown in their most prototypical color (Mohr, 2010; Wurm et al., 1993). All concepts were shown without background, centered on grey (5% black) background and in canonical view (Palmer, 1981). Size relationship was controlled within a category (as applied in the Snodgrass and Vanderwart image set, 1980) and small narrow objects were shown in 45° angle positions (Snodgrass & Vanderwart, 1980). Verbs were shown task-engaged (Thiessen et al., 2016). The graphic representations differed from the photographs in the sense that details that were not required for unambiguous recognition were simplified or omitted (e.g. labels on objects, patterns on garments, details of technical devices) (Biederman, 1987) (See Figure 2 as an example).

Figure 2: Depiction of the verb "diving". On the left as a photograph (*iStockPhoto.com, 487542208*, 2015), on the right as graphic representation.

Additionally, specific decisions were made to define the style of the graphic representation: No outline was used around the figures, therefore trying to avoid giving the image a schematic appearance and to differentiate the image style from the type of black-and-white line drawings. Textures were implemented in the form of different color shades and shapes (Brodie et al., 1991; Rossion & Pourtois, 2001), three-dimensionality was simulated by the use of shading and gloss spots (Adlington, 2009; Rossion & Pourtois, 2001). Further specific criteria were defined separately for the depiction of nouns and verbs (See Table 1 in Appendix: https://osf.io/kf7gv/?view_only=33c8687aa4e94907adec776d57c67922).

Image validation

To ensure the homogeneity of the stimuli used, the graphic representations were controlled for image agreement, visual complexity and name agreement prior to the main experiment in two related image validation studies. These variables were chosen because they have been shown to provide relevant information on the development of stimuli depicting concrete nouns (Fiez & Tranel, 1997; Snodgrass & Vanderwart, 1980). Sixty-two students from the FHNW Academy of Art and Design (23 males, 35 females, 4 without gender specification, $M_{age} = 25.6$, $SD = 6.3$) rated 128 graphic representations on 5-point scales for visual complexity and image agreement (1=very simple, 5=very complex; respectively 1=very low agreement, 5=very high agreement). Additionally, written name agreement was assessed for every term. The photographic stimuli were not included in the analysis as the results were expected to be similar to the graphic representations. The results of the pretest analysis were used to adapt the graphic representations not fulfilling the requirements before their use in the main study. (For a detailed description of this procedure and the adaptations see Table 2 and 3 in the Appendix).

Oral name agreement of the stimuli was examined in a second image validation study including 123 adults (50 males, 72 females, M_{age} = 41.6, SD = 18.5) and was assessed with the percentage of correctly answered items and the H-statistic (Snodgrass & Vanderwart, 1980). While higher H-statistic values indicate lower name agreement these scores were used to test the dominant term attributed to an image.

Sets

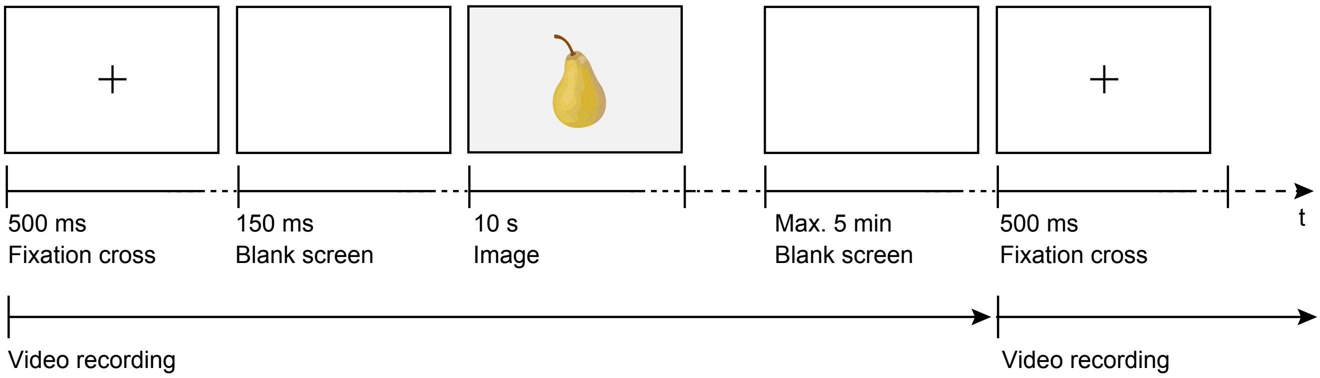
The photographs and graphic representations were compiled into eight different sets, in which the image type (photograph/graphic representation) and the word class (nouns/verbs) were controlled. Item listing was semantically and phonemically controlled for as well as for color to ensure maximal semantic, phonemic and/or illustrative distances between successive items (Mohr, 2010). Each set consisted of 64 image stimuli, half of them verbs, the other half nouns (Székely et al., 2005). Half of the terms were shown as graphic representations and the other half as photographs. Further, each set contained four trial stimuli. The sets were assigned randomly to the participants but ensuring that all sets were equally distributed among the participants.

Procedure

Participants with aphasia were tested at the facility of their language therapy. All other participants were tested at the FHNW University of Applied Sciences Northwestern Switzerland or in their private environment (due to the COVID-19 pandemic restrictions). The images were shown, and responses were video, and audio recorded on an Android Tablet via a mobile app developed for the study. The examiner was seated next to the participant (in 1.5m distance). Before starting the naming task of the main experiment, PWA completed two screening-tests: the Pyramids and Palm Tree Test (PPT; Howard & Patterson, 1992) and a shortened version of the Coloring of Pictures Test (CoPT; De Renzi et al., 1972). The CG did not complete the pre-tests.

The main experiment started with a standardized oral instruction of the naming task, followed by an example task. The introduction and the example task were repeated until the participant understood the task correctly. Each stimulus was presented with the App in the following manner: a fixation cross was presented for 500 milliseconds followed by a blank page for 150 milliseconds before the test item was presented for 10 seconds full screen on the tablet. After that a blank page was shown again for the maximum of five minutes, in order to provide a short break for the participant if needed (See Fig. 3). While the image was visible on the screen participants had to name the presented stimuli according to the beforehand given instructions: "Please name the image as fast and as correctly as possible by using one single word." After the participant had given a response or was clearly not able to provide one, the examiner clicked to show the next image. Besides the eight trial items, each participant named 64 images within four blocks of 16 images.

Figure 3: Example sequence of stimuli presentation with the mobile App.



Measures

Screening-Tests: Color Vision and Semantic Access

The Pyramid and Palm Tree Test (PPT; Howard & Patterson, 1992) and a shortened version of the Coloring of Pictures Test (CoPT; De Renzi et al., 1972) were conducted prior to the main experiment to measure the capacity to access semantic information about pictures (PPT) and unimpaired color vision (CoPT) in participants with aphasia. To pass the Pyramids and Palm Tree Test (PPT; Howard & Patterson, 1992) a minimum of 90% of the 47 items (adapted following Callahan et al., 2010; Mohr, 2010) had to be answered correctly. Coloring of Pictures Test (CoPT; De Renzi et al., 1972) was considered successful when three out of four colors were assigned correctly.

Naming correctness

Response of the naming task was considered correct if the participant provided exclusively the previously determined and therefore linguistically controlled reaction to the stimuli within 10 seconds. Only absolutely correct responses were included in the data analysis for this paper because the research question did not include a pragmatic perspective.

Naming latency

Naming latencies were only calculated for correct responses. Each naming latency was calculated as the interval between the onset of visual naming stimulus on the tablet screen and the onset of correct target response. Naming latencies were manually measured using the speech analysis program Praat (Boersma, 2001) which allows acoustic analysis by presenting waveform and spectrogram of audio data.

Invalid responses

Zero responses (no response or “I don’t know”) and technical errors, such as recording failure or loud noises, were excluded from the data set (Alario et al., 2004; Vorwerg et al., 2019).

Data analysis

The data was analyzed using generalized linear mixed and linear mixed models. These were estimated using maximum likelihood estimation offered by the statistical software R and the lme4-package (Bates et al., 2015). Tests were executed by applying the lmerTest package (Kuznetsova et al., 2017). The models were built in a forward selection approach. The best model fit was defined by running an F-Test on the Akaike information criterion (AIC) and Bayes information criterion (BIC).

Naming correctness

To analyze the data regarding the naming correctness, generalized linear mixed models with binomial distribution and logit link function were used. The final model contained the fixed effects of group (PWA vs. CG), image type (graphic representation vs. photograph), and word class (verb vs. noun). Participants were included as a random intercept.

Naming latencies

The naming latencies were analyzed with a linear mixed model with a Gaussian distribution and a link function. In the final model grouping variable (PWA vs. CG), image type (graphic representation vs. photograph), word class (verb vs. noun) were included as fixed effects without considering interactions, as well as a random intercept for each participant.

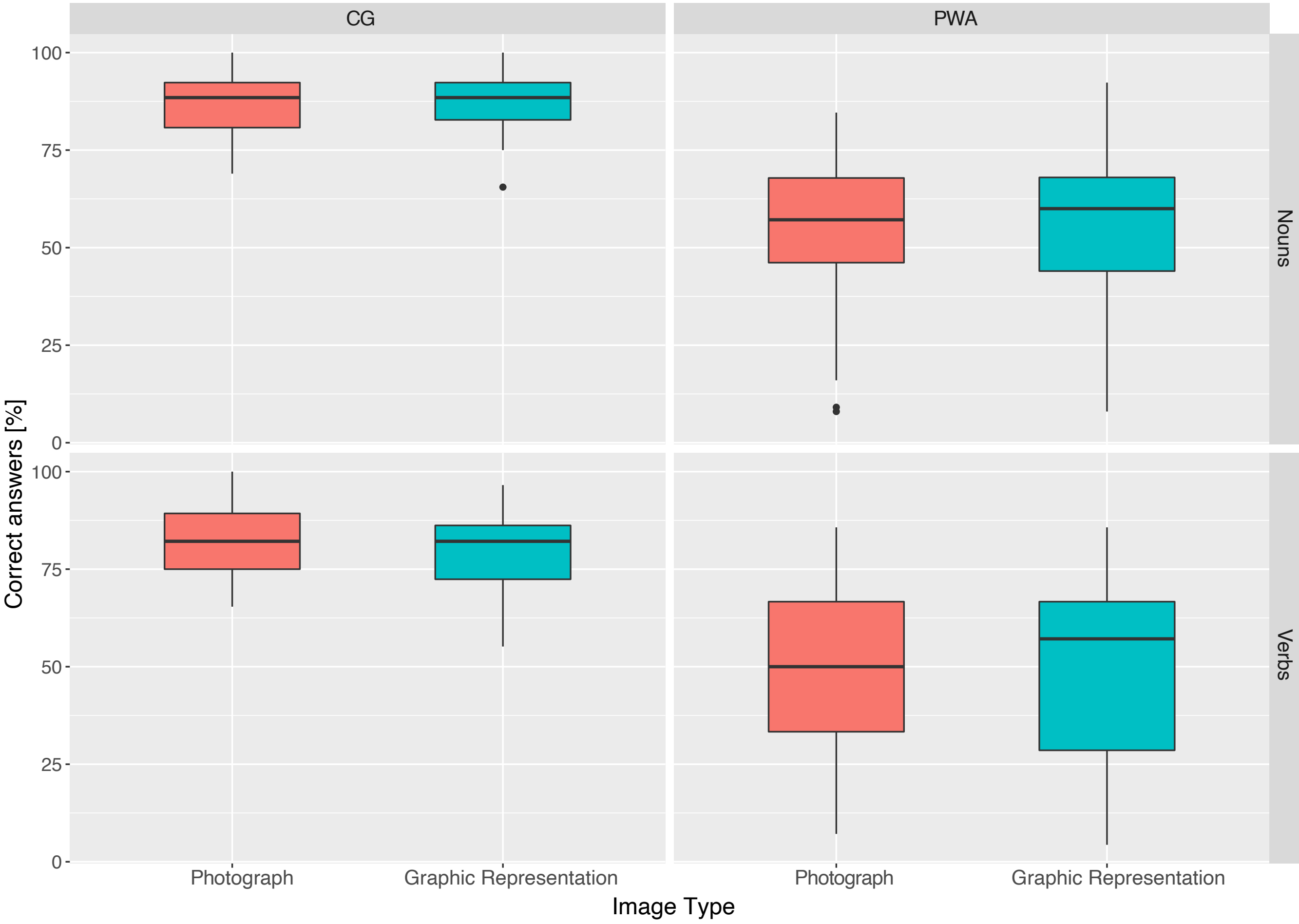
Results

Naming correctness

Participants in the CG named concepts significantly more correctly than PWA ($Z = -9.210$, $p < .001$) and nouns were named significantly more correctly than verbs ($Z = -5.712$, $p < .001$). The model revealed no significant main effect for the image type ($Z = -0.012$, $p = .967$). The best fitting model did not include any interaction effects. The CG named nouns depicted as photographs equally to graphic representation (Mean M ± Standard Deviation SD): $M_{Photo} = 86.0$, $SD = 8.4$ vs $M_{Graphic} = 87.0$, $SD = 7.5$. Verbs were also named equally, although less correct than nouns: $M_{Photo} = 82.3$, $SD = 9.4$ vs $M_{Graphic} = 79.5$, $SD = 10.6$.

In the PWA group, the mean of correct responses was slightly higher for naming nouns presented as graphic representation, although the model did not reveal a significant difference between the two image types (M±SD): $M_{Photo} = 53.7$, $SD = 21.0$ vs $M_{Graphic} = 56.0$, $SD = 18.5$. The mean for naming verbs correctly depicted as graphic representation was also higher than for naming verbs depicted as photographs: $M_{Photo} = 48.6$, $SD = 21.0$ vs $M_{Graphic} = 49.2$, $SD = 22.4$.

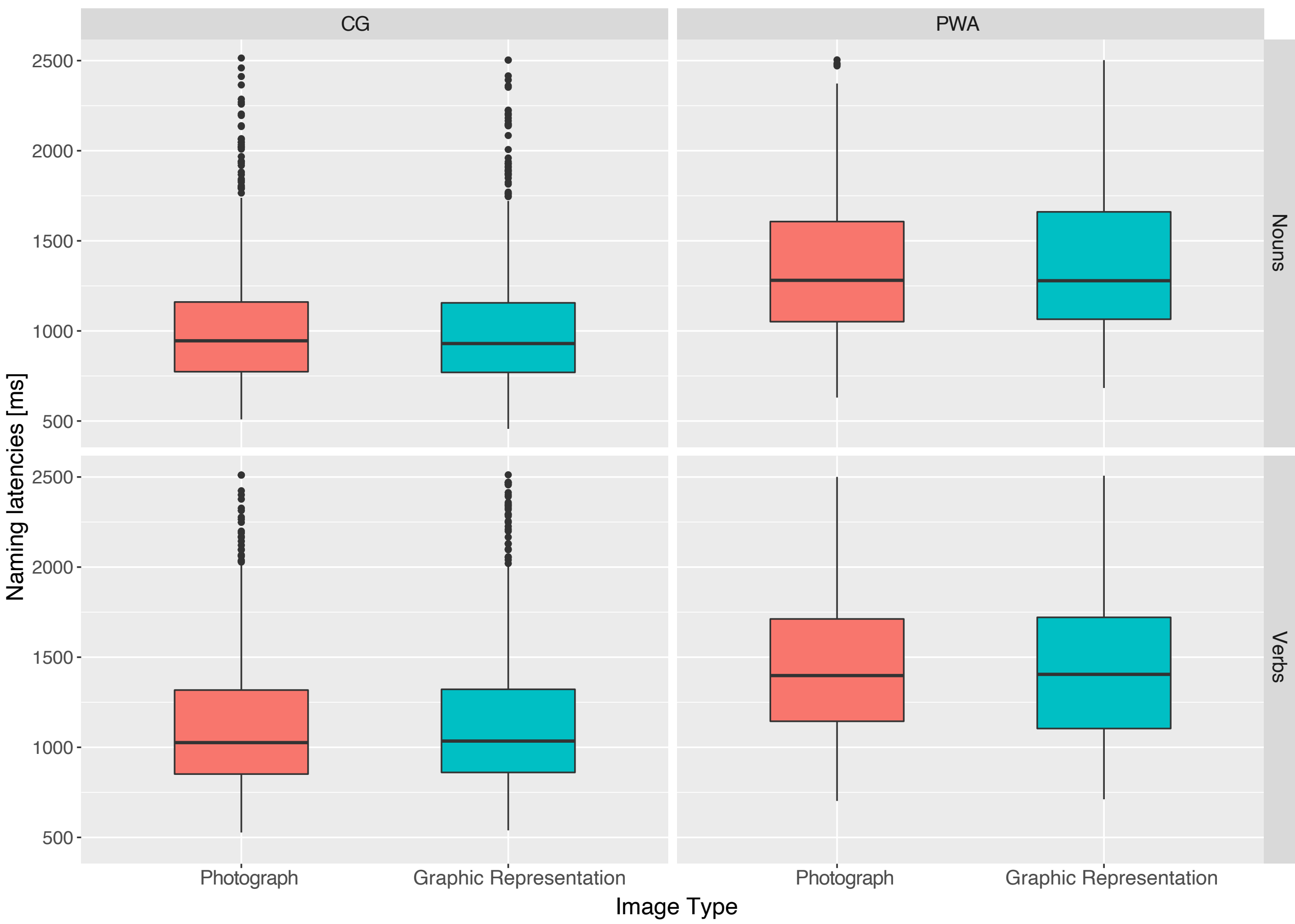
Figure 4: Naming correctness over all conditions for photographs and graphic representations. The box represents the interquartile range (IQR) ranging from the first to the third quartile and therefore the middle 50% of the data. The horizontal line indicates the median. Outliers are marked as dots and exceed the limit 1.5xIQR as defined by (Tukey, 1977).



Naming latencies

Regarding naming latencies, again two significant main effects were found, one for group (T = 8.333ms, p < .001) and one for word class (T = 10.689ms, p < .001): It took a participant significantly longer to answer if the stimulus was a verb. Also, as expected, PWA responded significantly slower than the CG. There was no significant effect for the fixed factor image type (T = 0.629ms, p = 0.529). In average the CG named images faster when naming nouns (M±SD): t_{Photo} = 1011.018ms ± 3337.184ms vs $t_{Graphic}$ = 1012.618ms ± 352.512ms, but slower in average when naming verbs: t_{Photo} = 1129.833ms, SD = 379.568ms vs $t_{Graphic}$ = 1137.495ms, SD = 396.958ms. PWA named nouns presented as photographs and graphic representations (t_{Photo} = 1356.504ms, SD = 396.433ms vs $t_{Graphic}$ = 1386.841ms, SD=426.573ms) faster than verbs (t_{Photo} = 1448.913ms, SD = 387.437ms vs $t_{Graphic}$ = 1455.634ms, SD = 431.588ms).

Figure 5: Naming latencies in milliseconds over all conditions for photographs and graphic representations. The box represents the interquartile range (IQR) ranging from the first to the third quartile and therefore the middle 50% of the data. The horizontal line indicates the median. Outliers are marked as dots and exceed the limit 1.5xIQR as defined by Tukey (1977).



Discussion

In this study a stimulus set of graphic representations depicting nouns and verbs was developed and correctness and naming latencies of both PWA and the CG were measured and compared to naming correctness and latencies of photographic stimuli. Naming correctness and naming latencies differed between the groups, with the CG naming images more correctly and faster than PWA. The two image types, however, did not have a significant effect on naming performance. This is in contrast to past studies that have shown effects of image type on naming performance and communication behavior (Bisiach, 1966; Ho et al., 2005; Ma et al., 2009). Specifically, when photographs have been compared to black-and-white line drawings results revealed an advantage for colored photographic stimuli over black-and-white line drawings (Griffith et al., 2014; Heuer, 2016). But these studies have compared image types that differ in relevant image features like color, texture and shading. Based on the results of our study, we showed that when drawn images are created with features that favor naming performance (e.g., Biederman & Cooper, 1992; Johnson et al., 1996; Naor-Raz et al., 2003; Palmer, 1981; Therriault et al., 2009) they can be named as correctly and as fast as photographs. Thus, graphic representations can be used alongside photographs in aphasia therapy and diagnosis, especially to represent concepts that cannot be easily photographed. In the production of new stimuli—especially verbs—, it remains to be decided whether it is more economical to find the appropriate image on a database, organize a photo shoot to depict people in complex activities such as “milk”, “fly” or “hunt”, or to engage a trained designer to depict the activities in the desired manner.

Our results further showed that word class affected naming in that sense, that nouns were named significantly more correctly and faster in the CG as well as in PWA. This is in line with previous studies, showing slower naming responses for verbs (Bastiaanse et al., 2003; Bastiaanse & Jonkers, 1998; Mätzig et al., 2009). We found no significant differences in naming performance of verbs in regard to image type. However, the results on naming correctness in PWA—although not on a significant level—showed that graphic representations were named more correctly on average than photographs. This could suggest that images, which are a reduced and translated version (described in our paper as graphic representations) of actions perceived in reality could be an advantage for depicting verbs compared to photographs which offer a more detailed representation. At the same time the response correctness in PWA varied more in naming graphic representations than in naming photographs of verbs which shows that there are greater interindividual differences in correctly naming graphic representations than in naming photographs of this word class. However, we could not conclusively demonstrate the advantages of graphic over photographic stimuli on naming performance of verbs and further research is needed to clarify the suitability of graphic representations as an image type for depicting verbs. Regardless of the question what image type is more suitable for naming verbs, the predominant use of line drawings in the image sets of verbs (Akinina et al., 2015; Khwaileh et al., 2018; Masterson & Druks, 1998; Székely et al., 2004) suggests that it might be easier to produce drawings for this word class than photographs. In our study, the large variance in naming an action correctly could be attributed to a varying degree of precision of the image itself, denoting the general difficulty to depict verbs in comparison to objects. This is also reflected in the underrepresentation of picture stimuli for verbs relative to those for objects (Akinina et al., 2015), though this gap has been addressed in the last two decades by the addition and elaboration of picture stimuli depicting verbs (Akinina et al., 2015; Fiez & Tranel, 1997; Masterson & Druks, 1998; Székely et al., 2004).

Limitations and future studies

In our study, we have shown that drawings (in our case graphic representations), when they include favorable image features like color, texture and shading, are named at least comparably to photographic stimuli. To further assess the image type of graphic representations for their suitability to depict nouns and especially verbs to be used in aphasia therapy and diagnostics, they would have to be compared with other image types, such as line drawings or pictograms. But more basically, there is still little knowledge about which image features (e.g., color, controlled size, shading) are specifically important for the representation of verbs. In our study, we have included favorable image features to the depictions of nouns and verbs equally. Therefore, it remains unclear which images features were especially helpful for naming graphic representations and which were relevant for naming photographs. Further studies thus could investigate image features separately regarding their impact on naming different word classes. Also, as it has been shown that image features for depicting nouns cannot fully be transferred to the depiction of verbs (Krull et al., 2003; Spezzano & Radanovic, 2010), more research on favorable image features for the representation of verbs is needed. Moreover, further studies may want to investigate which image types (e.g., photographs, graphic representations, line drawings) are best suited for what kind of word classes. Furthermore, naming correctness could also be dependent on the term depicted. In our study, we compared graphic representations to photographs depicting nouns and verbs. To specifically investigate interaction effects of depicted term, word class and image types, future studies could include additional word classes (e.g., adjectives, adverbs, non-representational nouns) and image styles (e.g., abstract patterns, pictograms).

Also, the group of PWA was very heterogeneous. Future studies could investigate whether certain image types are particularly suitable for a certain type of aphasia. In addition, it could possibly be that individuals, regardless of their language impairment, show a preference for an image type and that this could have an effect on naming performance. Further studies could investigate subjective influences and their impact not only on naming but also on the communication behavior of individuals (Griffith et al., 2014; Ho et al., 2005; Ma et al., 2009). Finally, we created an image set with limited scope. In order to use the set in aphasia therapy, it would need to be expanded to provide a larger number of stimuli. In particular, a greater ethnic and cultural diversity must be considered for the depicted persons engaged in an action in order to represent a larger population.

Conclusion

In aphasia practice, unambiguous images are needed for therapy and diagnosis. Unambiguous images are also required in research to study language processing in aphasia. Therefore, we support the call by Brown and Thiesssen (2017) for images that are meaningful and whose image features have been transparently studied for the use with PWA. We advocate for the creation of image stimuli that integrate image features that have a favorable influence on recognition and thus on naming performance in PWA. Although experienced designers produce representational drawings of high quality with relative ease (Kozbelt et al., 2010), the production of meaningful pictures and image material that is precisely suited to their intended use is time consuming and expensive. Alternatively, if photographs or other types of images are downloaded from the Internet, they should be carefully selected according to a coherent image concept that is appropriate for individuals with language deficits. Considering how much time and effort PWA invest in relearning and training language, it is essential to provide meaningful and unambiguous images that are produced or selected with great care.

Disclosure of Interest

This work was supported by the FHNW University of Applied Sciences and Art

Northwestern Switzerland as part of a Strategic Initiative 2018 - 2020. The authors report no conflict of interest.

Supplementary Material

The supplemental Material is available at:

https://osf.io/kf7gv/?view_only=33c8687aa4e94907adec776d57c67922

The image set is available under the licence of CC BY-NC-ND 4.0.

http://hdl.handle.net/20.500.11806/div/imageset_eInclusion_fhnw

References

Adlington, R. L. (2009). *The Influence of Surface Detail on Object Identification in Alzheimer's Patients and Healthy Participants* [PhD Thesis, University of Hertfordshire]. http://uhra.herts.ac.uk/handle/2299/4004

Akinina, Y., Malyutina, S., Ivanova, M., Iskra, E., Mannova, E., & Dragoy, O. (2015). Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*, *47*(3), 691–707. https://doi.org/10.3758/s13428-014-0492-9

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

APHAVOX. (o. J.). Abgerufen 28. April 2021, von https://aphavox.de/

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart Pictures: Effects of Age of Acquisition, Frequency, and Name Agreement. *The Quarterly Journal of Experimental Psychology Section A*, *50*(3), 560–585. https://doi.org/10.1080/783663595

Bastiaanse, R., Edwards, S., Mass, E., & Rispens, J. (2003). Assessing comprehension and production of verbs and sentences: The Verb and Sentence Test (VAST). *Aphasiology*, *17*(1), 49–73. https://doi.org/10.1080/729254890

Bastiaanse, R., & Jonkers, R. (1998). Verb retrieval in action naming and spontaneous speech in agrammatic and anomic aphasia. *Aphasiology*, *12*(11), 951–969. https://doi.org/10.1080/02687039808249463

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical SoftwareJournal of Statistical Software*, *67*(1), 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Ching Lu, C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. https://doi.org/10.3758/BF03196494

Benton, A. L., Smith, K. C., & Lang, M. (1972). Stimulus characteristics and object naming in aphasic patients. *Journal of Communication Disorders*, *5*(1), 19–24. https://doi.org/10.1016/0021-9924(72)90025-1

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147. https://doi.org/10.1037/0033-295X.94.2.115

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 121.

Bisiach, E. (1966). Perceptual Factors in the Pathogenesis of Anomia. *Cortex*, *2*(1), 90–95. https://doi.org/10.1016/S0010-9452(66)80030-8

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Brodie, E. E., Wallace, A. M., & Sharrat, B. (1991). Effect of surface characteristics and style of production on naming and verification of pictorial stimuli. *The American journal of psychology*, 517–545.

Brown, J., & Thiessen, A. (2018). Using images with individuals with aphasia: Current research and clinical trends. *American Journal of Speech-Language Pathology*, *27*, 504–515. https://doi.org/10.1044/2017_AJSLP-16-0190

Brumbi, J., Winkelmann, A., & Ablinger, I. (2017). *Einsatz der LetMeTalk-App bei globaler Aphasie* [Poster]. GAB.

Callahan, B. L., Macoir, J., Hudon, C., Bier, N., Chouinard, N., Cossette-Harvey, M., Daigle, N., Fradette, C., Gagnon, L., & Potvin, O. (2010). Normative data for the pyramids and palm trees test in the Quebec-French population. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *25*(3), 212–217. https://doi.org/10.1093/arclin/acq013

Corlew, M. M., & Nation, J. E. (1975). Characteristics of Visual Stimuli and Naming Performance in Aphasic Adults. *Cortex*, *11*(2), 186–191. https://doi.org/10.1016/S0010-9452(75)80043-8

Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arévalo, A., Cappa, S. F., & Padovani, A. (2007). Action and object naming in Parkinson's disease without dementia. *European Journal of Neurology*, *14*(6), 632–637. https://doi.org/10.1111/j.1468-1331.2007.01797.x

Cuetos, F., Gonzalez-Nosti, M., & Martínez, C. (2005). The picture-naming task in the analysis of cognitive deterioration in Alzheimer's disease. *Aphasiology*, *19*(6), 545–557. https://doi.org/10.1080/02687030544000010

Cutzu, F., & Edelman, S. (1994). *Canonical Views in Object Representation and Recognition* [Preprint]. http://cogprints.org/581/

De Bleser, R., & Kauschke, C. (2003). Acquisition and loss of nouns and verbs: Parallel or divergent patterns? *Journal of Neurolinguistics*, *16*(2), 213–229. https://doi.org/10.1016/S0911-6044(02)00015-5

De Renzi, E., Faglioni, P., Scotti, G., & Spinnler, H. (1972). Impairment in associating colour to form, concomitant with aphasia. *Brain*, *95*(2), 293–304. https://doi.org/10.1093/brain/95.2.293

Doak, C. C., Doak, L. G., & Root, J. H. (1996). Teaching Patients with Low Literacy Skills. *AJN The American Journal of Nursing*, *96*(12), 16M.

Fiez, J., & Tranel, D. (1997). Standardized stimuli and procedures for investigating the retrieval of lexical and conceptual knowledge for actions. *Memory & Cognition*, *25*(4), 543–569. https://doi.org/10.3758/BF03201129

Glaser, W. R. (1992). Picture naming. *Cognition*, *42*(1), 61–105. https://doi.org/10.1016/0010-0277(92)90040-O

Griffith, J., Dietz, A., & Weissling, K. (2014). Supporting Narrative Retells for People With Aphasia Using Augmentative and Alternative Communication: Photographs or Line Drawings? Text or No Text? *American Journal of Speech-Language Pathology*, *23*(2), S213–S224. https://doi.org/10.1044/2014_AJSLP-13-0089

Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology*, *30*(8), 943–961. https://doi.org/10.1080/02687038.2015.1081138

Ho, K. M., Weiss, S. J., Garrett, K. L., & Lloyd, L. L. (2005). *The Effect of Remnant and Pictographic Books on the Communicative Interaction of Individuals with Gl*. 16.

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test*. Pearson.

Howard, D., Patterson, K., Franklin, S., Orchard-lisle, V., & Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, *2*(1), 49–80. https://doi.org/10.1080/02643298508252861

Huber, W., Poeck, K., Weninger, D., & Willmes, K. (1983). *Aachener Aphasie Test*. Hogrefe.

Humphreys, G. W., Price, C. J., & Riddoch, M. J. (1999). From objects to names: A cognitive neuroscience approach. *Psychological Research*, *62*(2), 118–130. https://doi.org/10.1007/s004260050046

IStockPhoto.com, 487542208. (2015). https://www.istockphoto.com/de/foto/mann-taucher-tauchen-silhouette-isoliert-gm487542208-72997855

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive Components of Picture Naming. *Psychological Bulletin*, *120*(1), 113–139.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*.

Kemmerer, D. (2014). Word classes in the brain: Implications of linguistic typology for cognitive neuroscience. *Cortex*, *58*, 27–51. https://doi.org/10.1016/j.cortex.2014.05.004

Kim, M., & Thompson, C. K. (2000). Patterns of Comprehension and Production of Nouns and Verbs in Agrammatism: Implications for Lexical Organization. *Brain and Language*, *74*(1), 1–25. https://doi.org/10.1006/brln.2000.2315

Kohn, S. E., & Goodglass, H. (1985). Picture-naming in aphasia. *Brain and Language*, *24*(2), 266–283. https://doi.org/10.1016/0093-934X(85)90135-X

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language*, *91*(3), 294–302. https://doi.org/10.1016/j.bandl.2004.04.001

Kozbelt, A., Seidel, A., ElBassiouny, A., Mark, Y., & Owen, D. R. (2010). Visual selection contributes to artists' advantages in realistic drawing. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(2), 93–102. https://doi.org/10.1037/a0017657

Krull, R., Sharp, M., & Roy, D. (2003). Canonical views in procedural graphics. *IEEE International Professional Communication Conference, 2003. IPCC 2003.*

Proceedings., 22–28. https://doi.org/10.1109/IPCC.2003.1245462

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Laiacona, M., Luzzatti, C., Zonca, G., Guarnaschelli, C., & Capitani, E. (2001). Lexical and semantic factors influencing picture naming in aphasia. *Brain and Cognition*, *46*(1), 184–187. https://doi.org/10.1016/S0278-2626(01)80061-0

Laine, M., Vuorinen, E., & Rinne, J. O. (1997). Picture naming deficits in vascular dementia and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *19*(1), 126–140. https://doi.org/10.1080/01688639708403842

Leipzig Corpora Collection. (o. J.). *Leipzig Corpora Collection: Swiss German Wikipedia corpus based on material from 2012*. Dataset. https://corpora.uni-leipzig.de?corpusId=gsw_wikipedia_2012.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Ma, X., Boyd-Graber, J., Nikolova, S., & Cook, P. R. (2009). Speaking Through Pictures: Images vs. Icons. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, 163–170. https://doi.org/10.1145/1639642.1639672

Masterson, J., & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word frequency, familiarityand age-of-acquisition. *Journal of Neurolinguistics*, *11*(4), 331–354. https://doi.org/10.1016/S0911-6044(98)00023-2

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex*, *45*(6), 738–758.

https://doi.org/10.1016/j.cortex.2008.10.003

Mayer, G., & Villaire, M. (2007). *Health Literacy in Primary Care: A Clinician's Guide*. Springer Publishing Company.

Menn, L., & Bastiaanse, R. (2016). Beyond Chomsky versus Skinner: Frequency, language processing and aphasia. *Aphasiology*, *30*(11), 1169–1173. https://doi.org/10.1080/02687038.2016.1168920

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*(1), 131–147. https://doi.org/10.1016/S0749-596X(02)00509-0

Mohr, E. (2010). *Colour and Naming in Healthy and Aphasic People* [PhD Thesis, University of Durham, Department of Psychology]. https://core.ac.uk/display/85539/tab/similar-list

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is Color an Intrinsic Property of Object Representation? *Perception*, *32*(6), 667–680. https://doi.org/10.1068/p5050

Palmer, S. (1981). Canonical perspective and the perception of objects. *Attention and performance*, 135–151.

Pfab, J., Jakob, H., Spath, M., & Draxler, C. (2015). Neolexon—A Therapy App for Patients with Aphasia. *Interspeech*.

Readence, J. E., & Moore, D. W. (1981). A meta-analytic review of the effect of adjunct pictures on reading comprehension. *Psychology in the Schools*, *18*(2), 218–224. https://doi.org/10.1002/1520-6807(198104)18:2<218::AID-PITS2310180219>3.0.CO;2-1

Reymond, C., Müller, C., & Grumbinaite, I. (2019). E-Inclusion: Defining Basic Image Properties for Illustrated Stimuli in Aphasia Treatment. *Visible Language*, *53*(3), 4–27.

Richter, K., Wittler, M., & Hielscher-Fastabend, M. (2006). *Bielefelder Aphasie Screening (BiAS) – Zur Diagnostik akuter Aphasien*. https://pub.uni-bielefeld.de/record/1857986

Rossion, B., & Pourtois, G. (2001). Revisiting Snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, *1*(3), 413–413. https://doi.org/10.1167/1.3.413

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.

Spezzano, L. C., & Radanovic, M. (2010). Naming abilities: Differentiation between objects and verbs in aphasia. *Dementia & Neuropsychologia*, *4*(4), 287–292. https://doi.org/10.1590/S1980-57642010DN40400006

Stark, J. (1992). *ELA - Everyday Life Activities: Photo Series*. Jacqueline Stark.

Székely, A., Damico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arevalo, A., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25. https://doi.org/10.1016/S0010-9452(08)70174-6

Székely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250. https://doi.org/10.1016/j.jml.2004.03.002

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processing*, *10*(4), 335. https://doi.org/10.1007/s10339-009-0260-4

Thiessen, A., Beukelman, D., Hux, K., & Longenecker, M. (2016). A Comparison of the Visual Attention Patterns of People With Aphasia and Adults Without Neurological

Conditions for Camera-Engaged and Task-Engaged Visual Scenes. *Journal of Speech, Language & Hearing Research*, *59*(2), 290–301. https://doi.org/10.1044/2015_JSLHR-L-14-0115

Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., & Wise, R. (2006). The Role of Semantics and Grammatical Class in the Neural Representation of Words. *Cerebral Cortex*, *16*(12), 1790–1796. https://doi.org/10.1093/cercor/bhj115

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: Semantic representations of object nouns, action nouns and action verbs. *Journal of Neurolinguistics*, *15*(3), 317–351. https://doi.org/10.1016/S0911-6044(01)00037-9

Vorwerg, C., Suntharam, S., & Morand, M.-A. (2019). Language control and lexical access in diglossic speech production: Evidence from variety switching in speakers of Swiss German. *Journal of Memory and Language*, 40–53.

Widmer Beierlein, S., Reymond, C., Kuntner, K. P., Blechschmidt, A., Degen, M., Müller, C., Falcón García, N., Altermatt, S., Elsener, C., Karlin, S., Park, S., Reutimann, R., Parrillo, F., Rickert, E., Bucheli, S., Grumbinaite, I., Jochmann, A., Harvey, M., Lee, J., … Hemm, S. (2021). E-Inclusion – Ein interdisziplinäres, Schweizerisches Aphasie-Forschungsprojekt. *Aphasie und Verwandte Gebiete*.

Wilson, S., Eriksson, D., Schneck, S., & Lucanie, J. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PLOS ONE*, *13*, e0192773. https://doi.org/10.1371/journal.pone.0192773

Wurm, L. H., Legge, G. E., Isenberg, L. M., & Luebker, A. (1993). Color Improves Object Recognition in Normal and Low Vision. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 899–911.

Curriculum Vitae

Name: Claire Marie Sarah Reymond

Nationality: Swiss, Dutch

Date of Birth: December 3, 1973

Higher Education

2017 – 2022

Dissertation at the Human-Computer Interaction Research Lab,
Center for General Psychology and Methodology, Faculty of Psychology,
University of Basel, Switzerland

2010 – 2012

Master of Arts in Visual Communication and Iconic Research,
Academy of Art and Design, University of Applied Sciences Northwestern Switzerland

2005 – 2010

Bachelor of Science in Psychology, Faculty of Psychology,
University of Basel, Switzerland

1996 – 2000

Fachklasse für Grafik, The Basel School of Design, Basel, Switzerland

Work Experience

Since 2014

Scientific Associate, Institute Digital Communication Environments,
Academy of Art and Design, University of Applied Sciences Northwestern Switzerland

2019

Visiting Researcher, metaLAB and Berkman Klein Center, Harvard University, USA

2018 – 2021

Research Assistant Human-Computer Interaction Research Lab, Center for General
Psychology and Methodology, Faculty of Psychology, University of Basel, Switzerland

2013 – 2016

Lecturer Neue Schule für Gestaltung Langenthal, Switzerland

2003 – 2010

Independent Designer

Publications

2021

Vessel, E. A., Reymond, C., Etcoff N. (2021). Effect of Formal Training on Consensus of Visual Aesthetic Appreciation. Poster presentation at the 2021 Conference of the International Association of Empirical Aesthetics, London, UK.

2020

Reymond, C., Pelowski, M., Opwis, K., Takala, T., & Mekler, E. D. (2020). Aesthetic Evaluation of Digitally Reproduced Art Images. Frontiers in Psychology, 11.

https://doi.org/10.3389/fpsyg.2020.615575

Widmer Beierlein, S., Kuntner, K. P., Reymond, C., Degen, M., Shah, A., Müller, C., Falcón García, N., Karlin, S., Parrillo, F., Bucheli, S., Elsener, C., Reutimann, R., Park, S., Loew, J., & Hemm, S. (2020, Februar). E-Inclusion—Eine Benenn-App für Menschen mit Aphasie. 21. Jahrestagung des Deutschen Netzwerks für Evidenzbasierte Medizin, Basel. Schweiz.

Altermatt, S., Kuntner, K. P., Rickert, E., Wyss, S., Degen, M., Reymond, C., Widmer Beierlein, S., Blechschmidt, A., & Hemm, S. (2020). Automatic detection of naming latency from aphasia patients – using an extended threshold-based method. Conf Proc EMBEC Nov 2020, 71.

Hemm, S., Degen, M., Shah, A., Parrillo, F., Karlin, S., Altermatt, S., Blechschmidt, A., Bucheli, S., Kuntner, K. P., Widmer Beierlein, S., Reymond, C., Müller, C., & Grumbinaite, I. (2020). E-Inclusion—Eine Benenn-App für Menschen mit Aphasie. Projekt der Strategischen Initiative der HSL, PH und HGK, 2018–2020. Tag der Lehre der Pädagogischen Hochschule, Muttenz, Schweiz.

2019

Reymond, C., Müller, C., & Grumbinaite, I. (2019). E-Inclusion: Defining Basic Image Properties for Illustrated Stimuli in Aphasia Treatment. Visible Language, 53(3), 4.

Reymond, C., Müller, C., & Grumbinaite, I. (2019). Developing a New Set of 128 Images Illustrating Activities and Objects forTreatment Processes in Aphasia. Perception, 48, 1–236. https://doi.org/10.1177/0301006619863862

2017

Renner, M., Reymond, C., & Schubbach, A. (2017). Practice-led Iconic Research: Towards a Research Methodology for Visual Communication. Visible Language. The Journal of Visual Communication Research, 51(3-52.1), 1–208.

Reymond, C. (2017). Premises for Interaction between Images. Visible Language. The Journal of Visual Communication Research, 51(3-52.1), 148–173.

2013

Reymond, C. (2013). Bild und Bild: Eine Untersuchung zwischenbildlicher Prozesse. MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung, 23, 1–23. https://doi.org/10.21240/mpaed/23/2013.06.04.X

Reymond, C. (2013). Image and Image. An Investigation into Intericonic Processes. *36nd edition of the European Conference on Visual Perception 2013 Bremen*.