The Aesthetic Aha: On the pleasure of having insights into Gestalt

Claudia Muth *, Claus-Christian Carbon

Department of General Psychology and Methodology, University of Bamberg, Bamberg, Germany

ABSTRACT

Are challenging stimuli appreciated due to perceptual insights during elaboration? Drawing on the literature regarding aesthetic appreciation, several approaches can be identified. For instance, fluency of processing as well as perceptual challenge are supposed to increase appreciation: One group (Reber, Schwarz, & Winkielman, 2004) claims that fluency of processing increases appreciation. Others link aesthetics to engagement: Creation and manipulation of sense itself should be rewarding (Ramachandran & Hirstein, 1999). We experimentally tested the influence of insights during elaboration on liking. Pairs of stimuli – hardly detectable two-tone images including a face (Mooney face) and meaningless stimuli matched for complexity – were presented repeatedly. Having an insight as well as the intensity of the insight predicted subsequent gains in liking. This paper qualifies the role of insight (—aha!) on aesthetic appreciation through the effects of elaboration and problem-solving on understanding the processing of modern art.

1. Introduction

1.1. Appreciating difficult pictures: reward by fluency or challenge?

Why do we like perceptually challenging pictures? This ostensibly simple question is still yet to be answered: Fluency of processing as well as perceptual challenge are said to increase appreciation. Fluency theories assume that the more fluent the processing, the higher the appreciation (Reber et al., 2004). Evidence is provided by, e.g., the “mere exposure effect” (Zajonc, 1968), proposing an increase in preference with repeated, unreinforced exposure to stimuli. Also the preference for prototypes (Winkielman, Halberstadt, Fazendeiro, & Catty, 2006) and symmetric stimuli (Reber, 2002) is explained by fluency, as they are supposed to be easier to process than their opposites. These classical findings do, however, conflict with findings that associate novelty or innovativeness with high reward and liking (Blijlevens, Carbon, Mugge, & Schoormans, 2012; Carbon & Leder, 2005; Wittmann, Bunzeck, Dolan, & Düzel, 2007). Modern art also often impedes everyday perceptual routines while being popular at the same time. It offers various examples of perceptual challenge and sometimes sheer unresolvable contradictions (Meinhardt, 2009) and elicits “states of ambiguity, arousal, and uncertainty” (Jakesch & Leder, 2009, p. 2105) — like the football which is made of concrete in the artwork ‘jeu’ by Kristof Georgen. It produces a conflict between anticipated action and heavy material. Similar prediction errors were discussed and exemplified by Van de Cruys and Wagemans (2011), who claim that many artists combine familiar patterns with “a minimal deviation of default expectations” (p. 1043; see also the definition of indeterminacy by Pepperell, 2011, which “suggests the presence of objects but denies easy or immediate recognition”, p. 2). Also designers make use of visual–tactual incongruities to induce surprise in perceivers, which was found to augment a variety of emotions like interest, fascination, amusement, confusion, indignation and irritation (Ludden, Schifferstein, & Hekkert, 2012). In a similar fashion to the domain of music perception (Blood & Zatorre, 2001), these violations of expectation from visual cues might be linked to reward processing (Van de Cruys & Wagemans, 2011). The popularity of indeterminacy, surprise and contradiction in modern art and design obviously contradicts the often cited rule of ‘the easier the better’ once more and points to the necessity of incorporating further factors aside from fluency into research on aesthetic appreciation. Still, both ideas – that either easy or difficult stimuli increase appreciation – could have evolutionary advantages: Links between processing-fluency and reward could have been selected because familiarity signals harmlessness and fluency implies successful processing (Reber et al., 2004). Searching for novelty and challenges, on the other hand, might be rewarded in order to trigger exploration (Wittmann et al., 2007).
A third line of research proposes that neither easy nor difficult stimuli are preferred but that moderate amounts of ambiguity are maximally pleasurable (Jakesch & Leder, 2009). This is linked to the claim by Berlyne (1974) that the relationship between preference and arousal is described by an inverted U-shaped function. Too little arousal, as well as too much, decreases liking. When a stimulus is repeatedly presented, the increase in fluency thus would decrease arousal and increase liking as revealed by the “mere-exposure-effect”. Over-exposure after saturation, on the other hand, would lead to “under-arousal” and a decrease in liking. This limiting factor of boredom on the “mere exposure effect” is reflected by the discovery that complex objects increase the positive effect of exposure to a greater extent than simple ones (for the visual domain Bornstein, 1989; for the tactile domain: Jakesch & Carbon, 2012). Remarkably though, boredom was found not only to be associated with decreased (Pattyn, Neyt, Hendrickx, & Soetens, 2008), but also in some cases with increased, arousal (London, Schubert, & Washburn, 1972) (for an overview see Eastwood, Frischen, Fenske, & Smilek, 2012).

It is possible to combine typicality and novelty in music and design for the enhancement of pleasure (e.g., for consumer products see Hekkert, Snelders, & van Wieringen, 2003). This idea was recently re-investigated, stressing that arousal (by novelty) and familiarity (by prototypicality) both contribute, albeit independently, to aesthetic appreciation (Blijlevens et al., 2012). Such findings might explain the contradictory findings of preference for familiar stimuli (e.g. prototypes; Winkielman et al., 2006) and unfamiliar (or innovative) stimuli (Blijlevens et al., 2012; Carbon & Leder, 2005; Wittmann et al., 2007) discussed above. Nevertheless, a unified theoretical basis explaining the appeal of easy-to-process vs. difficult indeterminate stimuli is missing.

2. Methods

The major aim of the experiment was to test whether aesthetic appreciation benefits from insights during the elaboration of indeterminate stimuli. Two-tone images either containing a hidden Gestalt (i.e. a face) or not were repeatedly presented for half a second. After-insight moments of Gestalt detection were then related to the dynamics of liking ratings.

2.1. Preparation of material

Two pre-studies were conducted in order to evaluate and filter out a set of appropriate stimuli for the experiment. We used Face and NonFace stimuli. Pictures pertaining to the first category were based on photographs of faces taken from the website pixelio.de and the database of the Psychological Image Collection at Stirling (PICS). The original face was first blurred and then reduced to black and white so that recognition of the face was possible only after a period of elaboration (similar to so-called Mooney faces (1957); see Fig. 1).

Each of the Face pictures had a counterpart in the NonFace category that contained exactly the same elements arranged in a different non-facial composition by rotation and/or shift of parts of the face. In a first pre-study, six participants rated 98 stimuli (49 Face and NonFace, respectively) 11 times block-wise, after 500 ms of presentation on the question of whether they could detect a face in them, by pressing a key for either yes or no. Results showed that stimuli revealed faces too soon. Therefore the distance from the eyes to the screen was reduced from 40 to 30 cm in the experiment and random elements were added to the composition to make recognition harder. The face then appeared in the middle or at one corner of the picture so that the process of visual search was less efficient due to increased task demands. The possibility cannot be excluded that people interpreted unintended figural associations as faces. To reduce this risk, the composition of a stimulus was refined in cases when NonFace stimuli were reported to contain a face. Furthermore, an example of a face-pattern was given before the experiment. Instead of yes or no answers, the experiment used gradual scales for clearness of the face and similarity to a face to differ between recognition and guessing. Scaling also enabled a definition of ‘insight’ as the biggest difference between ratings for a stimulus in two succeeding blocks.
A further pre-study was conducted to exclude the possibility that a change in liking after such an insight was due to the attractiveness of the identified face. Eight participants rated the Mooney faces hidden in the Face stimuli on liking. Stimuli were excluded if a) they were liked significantly more or less than the empirical average liking rating and b) if participants noted verbally that they could not detect a face in one of them at all.

2.2. Experiment

2.2.1. Participants

Thirty participants aged 20 to 59 years (M_{age} = 29.5 years, SD = 9.4; 18 female, 12 male) volunteered for the experiment. Among them were seven university students of Psychology, nine university students of other disciplines, 11 workers, two high-school pupils, and one unemployed person. For their participation the Psychology students received course credits. All participants proved to have normal or corrected-to-normal vision through a standard Snellen Eye-chart test and normal colour vision assured by a short version of the Ishihara colour test consisting of four plates. None of them had participated in any of the pre-studies; they had no information about the aim of the study nor were they trained in art.

2.2.2. Apparatus and stimuli

The stimulus material consisted of 36 pictures, 18 belonging to the category Face or NonFace, respectively, selected on the basis of two pre-studies as described above. Participants were tested individually via PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) with an Apple Powerbook 17 in. with a resolution of 1440 × 900 pixels.

2.2.3. Procedure

A chin rest guaranteed that the distance between the eyes and the monitor was fixed at 30 cm, yielding an initial visual angle of the stimuli of 57.64° horizontally × 35.41° vertically and a final, minimal visual angle of 20.48° horizontally × 11.95° vertically. Participants were told that the aim of the study was to test the influence of presentation time on liking as a cover story. Stimuli were shown for 500 ms in a randomised order block-wise 13 times. The tasks alternated block-wise between choosing from a 7-point scale (1 = “not at all”, 7 = “very good”) how much one liked the picture and a detection block. The latter comprised two ratings on a 1 plus 7-point scale; first, how clearly one saw a face and in direct succession how similar it was to a human face (0 = “no face recognised”, 7 = “very clear” or “very”, respectively for similarity to a face; for an overview on the procedure see Fig. 2). The clearness and similarity tasks were explained after the first liking block to avoid the participants having searched for faces already during that first presentation. Additionally, an

![Fig. 1. Example of a pair of stimuli with (c) Face stimulus and (d) NonFace stimulus, plus the original photograph, the face was based on (a) and the mooneyised version of it (b).](image)

![Fig. 2. Procedure of the experiment: Stimuli were shown block-wise each for 500 ms. The task alternated between liking ratings and clearness/similarity ratings.](image)
example of a clear and face-like face was given before the second block. For the purpose of demonstration, elements that constitute the face were highlighted in red. The size of the pictures decreased every 2nd block by 20% (referring to the edge lengths) to make recognition easier within the course of the experiment as a result of more holistic processing [decreasing size is known to assist the recognition of visual closure and consequently Gestalt, see Gori and Spillmann (2010)].

2.2.4. Data analysis

We sought to analyse whether and how insights during elaboration influence aesthetic appreciation. Consequently, an analysis of changes in liking due to an insight was conducted. Insight was defined by the highest gain in clearness of, or similarity to, faces respectively, between two subsequent blocks per participant and stimulus (= maximum of all differences: block $n$ minus block $n-1$ per participant and stimulus). While in some cases clearness of, or similarity to, faces decreased and increased again in the course of elaboration, we chose only the first peak gain. That way the difference between the first sudden recognition of a Gestalt and the supposedly weaker experience of re- finding that pattern again was accounted for. We defined two types of insights: clearness insight for the highest clearness gain and similarity insight for the highest similarity gain. All liking ratings per participant and block were then shifted in regard to their temporal occurrence relative to this insight block: liking ratings directly after an insight formed one group, liking ratings in the subsequent liking block another, and analogously for all other pre- and post-insight liking ratings. This was done for clearness and similarity insights as well as for Face and NonFace stimuli separately (although we did not intend to induce an insight by NonFace stimuli, in some cases ratings of clearness or similarity were bigger than 0. These unexpected reports of increase in detection were consequently classified as insights in accordance with the described procedure). That way it was possible to compare liking directly before and after the insight block ("Δ to insight-block = 0"), liking ratings 2 blocks before the insight to those directly before the insight ("Δ to insight-block = −1") and so on adding up to 10 comparisons per stimulus category (Face or NonFace, respectively). Ratings were analysed by two-tailed paired t-tests, simple Regression Analyses and repeated measurement Analyses of Variance (ANOVA).

2.2.5. Results

Both types of insights had a major impact on the liking of Face stimuli, clearly demonstrated by the fact that liking only significantly increased directly after having an insight, revealed by two-tailed paired t-tests (significance levels were adjusted by using Bonferroni correction; significant p-values and effect sizes in Fig. 3a and b). Although NonFace stimuli also induced insights according to our definition, none of the differences between liking ratings before and after the insight was significant. Furthermore, the intensity of insights, defined as degrees of clearness or similarity ratings, showed direct influences on the degrees of liking. Simple regression analyses indicated explained variances of $R^2 = 0.685−0.946$ (i.e. $R^2$, see Table 1). Thus, liking increased in accordance with the insight's intensity.

Importantly, analyses of the development of liking over blocks revealed no evidence of increased liking over time and block progression as originally assumed by Zajonc (1968) and subsequent literature on the “mere exposure effect” (see Fig. 4): Repeated measurement Analyses of Variance (ANOVA) with a 2 [category: Face vs. NonFace] by 6 [block] factor design showed a main effect of category, $F(1,29) = 29.0, p < .0001, \eta^2 = .500$, with the category Face being more liked than NonFace (but liking of Face stimuli was not significantly higher during the first block, as a two-tailed paired t-test revealed, $M = .15, t = 1.6, p = .1369, n.s.$). The small but significant effect of block $F(6,174) = 2.8, p = .0117, \eta^2 = .089$, showed effects contrary to the typical results identified by “mere exposure effects”, because liking decreased over time. As the interaction failed to reach significance, $F(6,174) = 1.3, p = .2759, n.s.$, this decrease was found to be not category-specific.

Fig. 3. Changes in liking before and after the occurrence of an insight represented by effect sizes, Cohen’s $d$ of the mean difference between liking ratings directly before and after an insight (Δ to insight-block = 0), two blocks before and directly before an insight (Δ to insight-block = −1), directly after and two blocks after an insight (Δ to insight-block = +1) and so on for the following classes of insights: a) clearness insight and b) similarity insight. All significant effect sizes are highlighted by red diagonal stripes and added by the corresponding p-value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Conclusion

The present study revealed a major impact of insights into perceptual Gestalt on liking, which we will refer to as the “Aesthetic Aha effect” in the following. Participants who elaborated on indeterminate stimuli and gained insight into face-like appearance, showed strongly increased liking in a subsequent block of liking ratings. Importantly, during the whole course of the experiment we detected no other

<table>
<thead>
<tr>
<th>Face stimuli</th>
<th>NonFace stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>B</td>
</tr>
<tr>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Clearness on liking 8</td>
<td>.211</td>
</tr>
<tr>
<td>Similarity on liking 8</td>
<td>.303</td>
</tr>
</tbody>
</table>

Note: All reported regressions are significant at an alpha level of .01.
Our results stress that it is not the repeated presentation (i.e. the “mere exposure”) that increases liking when processing indeterminate stimuli. Rather the results point to the high relevance of the dynamics of elaboration and specifically of perceptual insight on the aesthetic process. We further revealed that such an Aesthetic Aha has quite a direct, but also temporarily limited effect on liking. Thus we propose that the assessment of perceptual and appreciative dynamics is indispensable for the understanding of the appreciation of difficult, indeterminate or otherwise challenging pictures as we find them in modern art (Jakesch & Leder, 2009). Our results are in accordance with the idea that the reduction of uncertainty is rewarding (Van de Cruys & Wagemans, 2011) and highlights the relevance of problem solving for aesthetic appreciation (Carbon & Jakesch, 2013; Dörner & Vehrs, 1975; Leder et al., 2004). At the same time, a transfer to the field of art perception demands clarification: does the Aesthetic Aha effect account for a) different kinds of aesthetic preference and b) for the variety of insights that can be induced during art perception?

4. Discussion

Our results stress that it is not the repeated presentation (i.e. the “mere exposure”) that increases liking when processing indeterminate stimuli. Rather the results point to the high relevance of the dynamics of elaboration and specifically of perceptual insight on the aesthetic process. We further revealed that such an Aesthetic Aha has quite a direct, but also temporarily limited effect on liking. Thus we propose that the assessment of perceptual and appreciative dynamics is indispensable for the understanding of the appreciation of difficult, indeterminate or otherwise challenging pictures as we find them in modern art (Jakesch & Leder, 2009). Our results are in accordance with the idea that the reduction of uncertainty is rewarding (Van de Cruys & Wagemans, 2011) and highlights the relevance of problem solving for aesthetic appreciation (Carbon & Jakesch, 2013; Dörner & Vehrs, 1975; Leder et al., 2004). At the same time, a transfer to the field of art perception demands clarification: does the Aesthetic Aha effect account for a) different kinds of aesthetic preference and b) for the variety of insights that can be induced during art perception?

a) In our study, we measured appreciation via explicit evaluations. As proposed by Makin, Pechinenda, and Bertamini (2012) these might be much more closely related to culture, experience and expectation than implicit automatic responses and therefore influenced additionally by a variety of factors besides fluency. This explains cases of misalignment between implicit and explicit preferences (Makin, Pechinenda, & Bertamini, 2012) and might also be important in regard to different modes of aesthetic processing governed by expertise. For instance, Cupchik (1995) proposes the distinction between reactive (relatively naïve perceivers’) and reflective (relative experts’) aesthetic processing. Whereas reactive processing is triggered by spontaneously pleasing or arousing features, for instance certain favourite colours or scenes but also familiar content, reflective processing is based on the active elaboration of a challenging stimulus and evokes more complex emotions. The latter consequently accounts better for modern artworks as they are defined mostly independent of superficial qualities and require deep elaboration before eliciting appreciation. The struggle to perceive a Gestalt in an indeterminate visual display might be special in this regard: beholders seem to refer the pleasure of their elaboration (having insights) to the stimulus’ explicit aesthetic appreciation. Along with the proposal by Makin, Pechinenda, and Bertamini (2012), the increase in fluency by insight might be even more relevant to implicit judgments which could be assessed via the Implicit Association Test as used by Makin, Pechinenda, and Bertamini (2012) or via the recently validated multidimensional IAT (mid-IAT Gattol, Ditye, Carbon, & Hutzler, 2007). Explicit judgments, on the other hand, might be influenced additionally by experience (Makin, Pechinenda, & Bertamini, 2012). We could thus use modern artworks instead of two-tone images in a follow up study and assess whether artistic expertise mediates the Aesthetic Aha effect on explicit judgments.

b) A second point concerns the problem-solving character of aesthetics. Our study predefined a problem (where is the face?) along with the according insight (detection of the facial Gestalt) quite explicitly. Elaboration of modern art, however, is mainly not a matter of easy processing and simple problem solving towards one specific pre-set solution. Instead, modern artworks often induce a firework of association dynamics, analogies, and transfers that serve to open up new levels of elaboration aside from the perceptual “core problem”. They force us to reflect on unsolvable ambiguities and even on perception itself (Meinhardt, 2009). Take the aforementioned example of Kristoff Georgen’s football made out of concrete. The reflexion on one’s own perception mechanism produces a valuable insight which however does not reduce the perceptual prediction error; the contradiction between felt affordance and given material. We thus suggest that insights can happen on various levels of elaboration, without necessarily leading to a unified “solution” or attribution of determinate meaning to the artwork (which might even offer unsolvable indeterminacy). Are there Cognitive Aha effects on appreciation analogous to the here reported Aesthetic Aha effect? Can we speak of problem-solving in terms of the sudden emergence of a Gestalt at all? This point exposes a methodological issue: to evoke and control for an insight, it is necessary to pose a task to participants and thus induce a search for a solution which might hinder free and spontaneous elaboration. Furthermore, Makin, Wilton, Pechinenda, and Bertamini (2012) found that people show positive affective responses to symmetry (by smiling) if they are asked to look for it — interestingly, the same accounts for randomness. In accord with our findings, it might thus not only be the features of an object (e.g. symmetry) or of processing (fluency) but the successful discovery of a target pattern itself that influences appreciation.

Our results reveal that perceptual insights into Gestalt within difficult pictures increase appreciation — an effect which might play an important role in the appreciation of modern art. Nevertheless, the spontaneity and variety of possible insights during art perception pose challenges for empirical research. A phenomenological approach
might find various levels of ambiguity and insight aside from Gestalt-detection bringing us closer to the multisided nature of aesthetic experience — and the on-going fascination for it.

Acknowledgements

We would like to thank Claudia Gorr and Stefanie Wolz for text revision and Alun Brown for proofreading. Thanks also for very valuable comments and insights (!) provided by Alexis Makin, Johan Wagemans, Marco Bertamini and Sander Van de Cruys.

References


