

# The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of Abstract Art

Evidence for a Default–interventionist Account

Ball, Linden J.; Threadgold, Emma; Marsh, John E.; Christensen, Bo T.

*Document Version*

Accepted author manuscript

*Published in:*

Metaphor and Symbol

*DOI:*

[10.1080/10926488.2018.1481255](https://doi.org/10.1080/10926488.2018.1481255)

*Publication date:*

2018

*License*

Unspecified

*Citation for published version (APA):*

Ball, L. J., Threadgold, E., Marsh, J. E., & Christensen, B. T. (2018). The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of Abstract Art: Evidence for a Default–interventionist Account. *Metaphor and Symbol*, 33(3), 235-252. <https://doi.org/10.1080/10926488.2018.1481255>

[Link to publication in CBS Research Portal](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Take down policy

If you believe that this document breaches copyright please contact us ([research.lib@cbs.dk](mailto:research.lib@cbs.dk)) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025



# The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of Abstract Art: Evidence for a Default-interventionist Account

**Linden J. Ball, Emma Threadgold, John E. Marsh, and Bo T. Christensen**

Journal article (Accepted manuscript\*)

**Please cite this article as:**

Ball, L. J., Threadgold, E., Marsh, J. E., & Christensen, B. T. (2018). The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of Abstract Art: Evidence for a Default-interventionist Account. *Metaphor and Symbol*, 3(3), 235-252. <https://doi.org/10.1080/10926488.2018.1481255>

This is an Accepted Manuscript of an article published by Taylor & Francis in *Metaphor and Symbol* on 01 Aug 2018, available online:

DOI: <http://www.tandfonline.com/10.1080/10926488.2018.1481255>

\* This version of the article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the publisher's final version AKA Version of Record.

Uploaded to [CBS Research Portal](#): August 2019



**This is the final manuscript version of an article that has been accepted for publication in *Metaphor and Symbol*. The published version of record may differ from this manuscript.**

**The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of  
Abstract Art: Evidence for a Default–Interventionist Account**

Linden J. Ball<sup>1</sup>, Emma Threadgold<sup>1</sup>, John E. Marsh<sup>1,2</sup>,

and

Bo T. Christensen<sup>3</sup>

<sup>1</sup>School of Psychology, University of Central Lancashire, Preston, UK

<sup>2</sup>Department of Building, Energy & Environmental Engineering, University of Gävle, Gävle,  
Sweden

<sup>3</sup>Department of Marketing, Copenhagen Business School, Copenhagen, Denmark

Running Head: AESTHETIC JUDGMENTS OF ABSTRACT ART

\*Corresponding Author:

Linden J. Ball

School of Psychology

University of Central Lancashire

Preston, PR1 2HE, UK

Tel: +44 (0)1772 893421

Email: LBall@uclan.ac.uk

## Abstract

We report an experiment investigating how stimulus complexity and conceptual fluency (i.e., the ease of deriving meaning) influence aesthetic liking judgments for abstract artworks. We presented participants with paintings at two levels of complexity (high vs. low) and five levels of conceptual fluency (determined from a prior norming study) and requested separate ratings of beauty and creativity. Our predictions were derived from the PIA Model (Pleasure-Interest Model of Aesthetic Liking by Graf & Landwehr, 2015), which views aesthetic preferences as being formed by two, distinct fluency-based processes: an initial, automatic, stimulus-driven, *default* process and a subsequent, perceiver-driven *deliberative* process. A key trigger for deliberative processing is assumed to be *disfluency* at the default stage, as caused by factors such as visual complexity. We predicted that complexity and conceptual fluency would interact in determining aesthetic liking, with people preferring complex stimuli, but only when these are relatively easy to process conceptually. Our results supported this prediction for beauty judgments, although creativity judgments showed a curiously uniform profile. Nevertheless, the predictive capacity of the PIA Model in relation to beauty judgments attests to the explanatory strength of this *default-interventionist* theory of aesthetic liking. We conclude by noting important parallels between the PIA Model and the Revised Optimal Innovation Hypothesis (Giora et al., 2017), which likewise has broad reach in explaining how defaultness and non-defaultness affect pleasure across a range of linguistic and pictorial stimuli.

# **The Effects of Stimulus Complexity and Conceptual Fluency on Aesthetic Judgments of Abstract Art: Evidence for a Default–Interventionist Account**

The study of aesthetics and its causal determinants is one of the oldest disciplines of experimental psychology, deriving from the foundational research of Fechner (1876), who aimed to draw links between the objective properties of stimuli and people’s aesthetic responses such as their judgments of beauty. This theme was taken up in earnest in the 20<sup>th</sup> century by Berlyne (e.g., 1971, 1974), who focused on the underlying physiological mechanisms mediating between objective stimulus properties and aesthetic responses, with an emphasis on arousal states. More recently, Berlyne’s insights and empirical findings have been refined as part of the contemporary “processing-fluency approach” to explaining aesthetic pleasure (for a review see Reber, Schwarz, & Winkielman, 2004a), which claims that people’s aesthetic judgments are closely aligned with the subjective ease with which mental operations are performed when perceiving an object.

Reber et al.’s (e.g., 2004a) processing-fluency theory has had a major bearing on the study of aesthetics and likewise has much relevance to the experiment we report here, which was concerned with the relation between processing fluency and aesthetic liking in the context of aesthetic judgments about abstract artworks. In motivating our study, we were keen to focus on a fluency factor, that of *conceptual fluency*, which has received relatively little attention in the literature on empirical aesthetics. Conceptual fluency is associated with the ease of deriving *meaning* from stimuli (Whittlesea, 1993; Topolinski & Strack, 2009). Our interest in this factor contrasts with most extant research on aesthetics, which has instead focused on fluency experiences arising from the perceptual properties of stimuli (e.g., their symmetry). Such perceptual factors give rise to so-called *perceptual fluency* effects on aesthetic preferences, which we discuss more fully below. Some research has also examined *retrieval fluency* (e.g., Winkielman, Schwarz, & Belli, 1998), which is the ease with which

previously encountered stimuli can be brought to mind, with repeated prior exposure being one well-established variable known to make stimuli easier to retrieve and process (Zajonc, 1968). We contend, however, that the fluent extraction of conceptual meaning from stimuli – especially in the context of abstract art appreciation – should be at least as important to judgments of aesthetic preference as more basic perceptual fluency effects.

Our study also manipulated a second factor that was orthogonal to conceptual fluency, that of *stimulus complexity*, since this factor is of increasing interest to researchers attempting to understand the basis of aesthetic judgments (e.g., see Landwehr, Labroo, & Herrmann, 2011) and may be especially pertinent in the context of appreciating abstract artworks, which seem to range widely in their perceived complexity. We additionally note that in directing our research at the issue of how conceptual fluency and stimulus complexity impact aesthetic judgments, we wished to examine how these factors might influence two specific types of aesthetic preference, that is, judgments of *beauty* and judgments of *creativity*. Although there is a large literature on how fluency is related to beauty judgments, there is limited research on how fluency impacts upon creativity judgments, despite creativity seeming to be a similarly important aesthetic dimension when evaluating artworks.

Before presenting more details of our study and the theoretical ideas driving our predictions, it is important to mention that a new conceptual approach has recently been posited that both challenges and extends Reber et al.'s (e.g., 2004a) dominant processing fluency approach to aesthetic preference. This new approach, advanced by Graf and Landwehr (2015), presents a *dual-process* perspective on issues relating to fluency-based aesthetics in an effort to reconcile inconsistent findings in the literature regarding aesthetic preference patterns. Graf and Landwehr's theory proposes that aesthetic preferences can be formed as a result of two, distinct fluency-based processes that are organised sequentially, that is: (i) an immediate, *default*, stimulus-driven process that occurs upon encountering an

aesthetic object, which gives rise to initial aesthetic judgments of pleasure or displeasure; and (ii) a subsequent, perceiver-driven, *deliberative* process arising from a person's motivation to process a stimulus further, which engenders fluency-based aesthetic evaluations of interest, boredom or confusion. As we explain below, we find aspects of Graf and Landwehr's (2015) dual-process theory to be very compelling in the context of predicting people's aesthetic judgments about abstract artworks under manipulations of conceptual fluency and complexity, not least because the theory seems to be able to capture the interplay between initial, stimulus-driven judgments and perceiver-driven override of such default evaluations.

### **The Processing-Fluency Account of Aesthetic Liking**

As we have noted, the dominant contemporary approach to understanding the mechanisms underlying aesthetic preference is Reber et al.'s (e.g., 2004a) processing-fluency account, which involves three foundational assumptions (see Graf & Landwehr, 2015). First, dependent upon an object's perceptual properties and a perceiver's prior processing experience with the object, the processing of the object will be experienced as more or less fluent. Second, the experience of fluent processing feels subjectively good at an affective level. Third, if this positive affect is not attributable to a different source, then it directly drives the aesthetic evaluation of the object, leading to a strong *fluency–liking* link whereby perceivers demonstrate a greater aesthetic preference for objects that are fluently processed compared to those that are disfluently processed (Reber et al., 2004a; Reber, Winkielman, & Schwarz, 1998; Winkielman, Halberstadt, Fazendeiro, & Catty, 2006).

In support of this predicted fluency–liking link is abundant evidence that objective stimulus properties that facilitate fluent processing (e.g., symmetry, visual clarity, figure-ground contrast, exposure duration and curvature of contours) also increase aesthetic liking (e.g., Bar & Neta, 2006; Bertamini, Makin, & Rampone, 2013; Humphrey, 1997; Oppenheimer & Frank, 2008; Reber, 2002; Reber et al., 1998; Reber, Wurtz, &

Zimmermann, 2004b; Wurtz, Reber, & Zimmermann, 2008). Likewise, factors that are linked to a perceiver's prior history with a stimulus can increase processing fluency during either retrieval or recognition and correspondingly increase aesthetic liking, with key factors being repeated exposure (e.g., Bornstein, 1989; Bornstein & D'Agostino, 1994) and perceptual priming (e.g., Reber et al., 1998).

As has been pointed out recently by Graf and Landwehr (2015), however, despite the broad-ranging evidence for a fluency–liking link, there has been a growing body of findings that challenge the assumptions of the processing-fluency approach to aesthetic preference, and particularly its emphasis on the notion that aesthetic judgments arise primarily as a function of an object's perceptual properties and a perceiver's prior processing experience with the object. Graf and Landwehr note two main sources of evidence that challenge the fluency-based account. First, several studies have revealed the existence of an inverted U-shaped relationship between processing fluency and aesthetic preference (e.g., Berlyne, 1970; Hekkert, Snelders, & van Wieringen, 2003; Landwehr et al., 2013). This evidence clearly runs counter to the assumption that there should be a monotonically increasing relationship between processing fluency and aesthetic liking. Second, some studies have observed a positive effect of *novelty* on aesthetic liking (e.g., Blijlevens, Carbon, Mugge, & Schoormans, 2012; Giora, Fein, Kotler, & Shuval, 2015; Giora, Fein, Kronrod, Elnatan, Shuval, & Zur, 2004; Giora, Givoni, Heruti, and Fein, 2017; Hekkert et al., 2003) as well as a positive effect of *visual complexity* (e.g., Landwehr et al., 2011; Martindale, Moore, & Borkum, 1990). From the perspective of the processing-fluency framework, it would be expected that novelty and complexity should decrease processing fluency and thereby decrease aesthetic liking. This inverse complexity effect is particularly fascinating and further underscores why we deemed it valuable in the present study to explore complexity as a factor that might be

associated with aesthetic judgments of abstract artworks, which are notorious for often being complex in nature.

### **The Dual-Process Approach to Aesthetic Liking**

Graf and Landwehr's (2015) dual-process approach to understanding fluency-based aesthetics has arisen as a direct response to increasing evidence of the kind just highlighted, which is difficult to explain according to a standard processing-fluency account. In developing their framework, Graf and Landwehr are particularly persuaded by theorists who argue that a perceiver's active *cognitive elaboration* of a stimulus can play a key role in aesthetic liking (e.g., Armstrong & Detweiler-Bedell, 2008; Carbon & Leder, 2005; Muth & Carbon, 2013; Muth, Pepperell, & Carbon, 2013). In other words, a perceiver is proposed not always merely to react passively to a stimulus, but is also assumed to have the potential to engage actively with a stimulus, such as arises when they devote additional processing toward gaining a deeper interpretation and understanding of it.

There is, indeed, some good empirical evidence that supports the importance of active cognitive elaboration in contributing to aesthetic liking. For example, studies using paintings as stimuli that have manipulated the presence of titles (e.g., Leder, Carbon, & Ripsas, 2006; Millis, 2001) and/or accompanying descriptive or stylistic information (e.g., Belke, Leder, & Augustin, 2006; Russell, 2003), have revealed that the aesthetic appreciation of such artworks can be enhanced through elaborative processing – provided that such elaboration is associated with a *meaningful* analysis of the stimulus. Likewise, active processing that is cued by instructing participants to evaluate stimuli on several dimensions can positively influence their appreciation of product designs (e.g., car exteriors), but only when such designs are novel, innovative or atypical (Carbon, Faerber, Gerger, Forster, & Leder, 2013; Carbon & Leder, 2005; Faerber, Leder, Gerger, & Carbon, 2010; Landwehr et al., 2013).

Graf and Landwehr (2015) propose that this latter research provides compelling evidence that there can be a positive relationship between elaborative processing and aesthetic liking in situations where the stimulus holds an “appropriate elaboration affordance”, which often arises when initial *disfluent* processing has occurred. Moreover, they propose that inconsistencies in the literature on empirical aesthetics relating to the observed association between processing fluency and liking can be directly attributed to the failure of the standard processing-fluency account to embrace fully the possibility that perceivers can take an active role in processing an initially disfluent stimulus so as to overturn early feelings of disliking and replace them with feelings of liking. The way in which meaningful elaborative processing can potentially modulate initial processing disfluency so as to enhance positive, aesthetic judgments formed a key driver for our study.

In an attempt to provide a comprehensive theoretical framework that can explain apparent empirical contradictions arising from studies of aesthetic judgments, Graf and Landwehr (2015) advance their “Pleasure Interest Model of Aesthetic Liking” (PIA Model) as a form of *default–interventionist* dual-process theory (cf. Evans, 2017; Evans & Stanovich, 2013; Kahneman, 2011). The PIA Model is centrally concerned with capturing the idea that aesthetic preferences can be formed as a result of two, qualitatively different and sequentially organised processes. More specifically, aesthetic liking deriving from the first process is based on stimulus-driven cognitive operations that occur mandatorily upon encountering an aesthetic object and which preclude deeper elaboration because of their unintended and automatic nature. In essence this first process provides a *default* aesthetic response of initial pleasure or displeasure. In contrast, aesthetic liking, arising from the second process, reflects an interaction between the stimulus-based output and perceiver-driven elaboration that can give rise to a more reflective and deliberative aesthetic evaluation that may be positively or negatively valenced. Graf and Landwehr conceptualise this second process as involving

“higher-order” cognitive operations such as careful and deliberate stimulus analysis as well as meaning assignment and interpretation.

We note the close alignment between Graf and Landwehr’s default–interventionist dual-process theory of aesthetic liking and recent research on “meta-reasoning” in the literature on judgment, decision-making and reasoning (e.g., Ackerman & Thompson, 2017a, 2017b). Such meta-reasoning research likewise adopts a dual-process stance on human thinking, emphasising how metacognitive monitoring is based on a variety of cues, including ones that derive from the perceivable features of the task (e.g., its apparent complexity) as well as from one’s own experience of attempting the task (e.g., subjective ease of processing). Ackerman and Thompson’s (2017a, 2017b) “meta-reasoning framework” – like Graf and Landwehr’s PIA Model – similarly entertains the possibility that initial *disfluent* processing in reasoning contexts may trigger more deliberative, analytic processing (see Alter, Oppenheimer, Epley, & Eyre, 2007; Thompson et al., 2013a, 2013b; but see also Ball, Threadgold, Solowiej, & Marsh, in press; Meyer et al., 2015). As such, Graf and Landwehr’s (2015) PIA Model complements major developments in the fields of judgment, decision-making and reasoning, which we would argue here is a factor that adds to its credibility.

That said, there are many important theoretical ideas and subtle distinctions that underpin Graf and Landwehr’s theorising (see also Graf & Landwehr, 2017), which also render it somewhat distinct from other default–interventionist theories as well as the meta-reasoning framework. Although we do not wish to get sidetracked either into reviewing the full model or examining recent debates regarding its potential limitations (e.g., Consoli, 2017), we nevertheless wish to emphasise three key assumptions of the model that relate closely to the factors that we manipulated in the study we report below. We therefore rehearse these assumptions here in order to show how they directly informed our predictions.

The first assumption to mention is that although Graf and Landwehr acknowledge that initial, default processing *and* subsequent, deliberative processing can both be potentially influenced either by perceptual fluency or by conceptual fluency, they nevertheless propose that default processing is relatively more influenced by perceptual fluency and that deliberative processing is relatively more influenced by conceptual fluency. This is because perceptual fluency is a largely passive, stimulus-driven experience, whereas conceptual fluency is a largely active, perceiver-driven process that places a heavy burden on elaborative and interpretative processes. In the context of our study, this first assumption implies: (i) that complexity (i.e., a stimulus-based property) will primarily have an impact at the default processing stage, with higher complexity promoting an increased sense of disfluency; and (ii) that conceptual fluency (arising from the ease of meaning extraction and stimulus interpretation) will primarily have an impact at the deliberative processing stage.

The second assumption of the PIA Model to note is that the cue to move from default processing to deliberative processing is considered to be jointly determined by an interplay between two factors: (i) a feeling of *disfluency*, which signals to the perceiver the need to invest more effort in processing the stimulus; and (ii) the perceiver's need for cognitive enrichment. Thus, for example, if a person experiences disfluency during default processing and has a high need for cognitive enrichment then the motivation to engage in deliberative processing will be especially strong. In cases where disfluency and need for cognitive enrichment are in opposition then it is the *relative strength* of these factors that determines whether deliberative processing is triggered. In our study we would expect that increasing levels of disfluency (arising from stimulus complexity) would be likely to trigger deliberative processing of presented stimuli even when balanced across individual differences in need for cognitive enrichment, which was not measured as part of our procedure.

The third assumption of the PIA Model that is of direct relevance to our study concerns the way in which deliberative processing can lead to aesthetic evaluations. What is initially perceived to be a disfluent stimulus through default processing may subsequently be found to be relatively easy to integrate into existing knowledge structures when processed deliberately. The updated fluency level that is experienced after deliberative processing can thereby lead to a final aesthetic evaluation that is markedly different to the initial aesthetic evaluation that arose from default processing. Indeed, of particular applicability to our study is this very case where initial perceptions of *disfluency* arising from stimulus complexity give rise to deliberative processing that is considered to be *less difficult than expected* because of the relative ease of extracting meaning from the stimulus as a consequence of its high conceptual fluency. In this situation the conceptually fluent deliberative processing that arises should, according to Graf and Landwehr, result in a positive aesthetic judgment of liking for what was initially a relatively disfluent stimulus.

### **Predictions**

In relation to the present study, which simultaneously manipulated the complexity (high vs. low) of presented abstract artworks and their conceptual fluency (across five linearly increasing levels), the aforementioned assumptions of the PIA Model give rise to two essential predictions. The first is that there should be a main effect of conceptual fluency on aesthetic liking such that conceptually fluent artworks will be liked to a greater degree than conceptually disfluent artworks. This prediction reflects the assumption that being more readily able to derive a meaningful interpretation of an abstract artwork at either the default processing stage or the deliberative processing stage should be a relatively pleasurable experience, especially if in general terms perceivers approach abstract artworks with an initial presupposition that they will be relatively difficult to process meaningfully.

The second prediction is that the main effect of conceptual fluency on aesthetic liking should be modulated by stimulus complexity, giving rise to a conceptual fluency by complexity interaction. This interaction is a consequence of the PIA Model's assumption that it is *complex* stimuli (i.e., those that are relatively disfluent at the default processing level) that will trigger more deliberative processing and effort-after-meaning rather than *simple* stimuli (i.e., those that are relatively fluent at the default processing level). The result of complex stimuli being subjected to such deliberative processing is that they will be associated with increased aesthetic liking compared to simpler stimuli, with this increased liking being especially marked for stimuli that are also more conceptually fluent. In other words, people should tend to like best of all the abstract artworks in our study that initially seem to be complex but which turn out to be relatively easy to derive meaning from. In contrast, abstract artworks that initially seem to be complex and which then remain hard to derive meaning from will persist in being fairly unappealing, much like their simpler counterparts.

In formulating our predictions, we leave open the possibility of observing a main effect of stimulus complexity on aesthetic liking. We note that given the nature of the predicted interaction, it is entirely possible that stimulus complexity could emerge as a significant main effect in its own right if judgments of aesthetic liking show a large influence of complexity at the higher levels of conceptual fluency. That said, the critical aspect of our predictions is the expected interaction between conceptual fluency and stimulus complexity, since this interaction derives very specifically from the PIA Model.

In terms of dependent measures, we examined two dimensions of aesthetic liking, that is, beauty and creativity. Although there is a fairly extensive literature on how fluency is related to judgments of beauty, there is contrastingly very little research on how fluency influences judgments of creativity. We also note that the PIA Model does not address creativity as a form of aesthetic liking, focusing instead on pleasure/displeasure and the

dimensions of interest, boredom and confusion. Perceived beauty and creativity, however, do seem to be subjective experiences that relate centrally to aesthetic appreciation such that they may well be impacted in similar ways by manipulations of stimulus complexity and conceptual fluency. That is, creativity evaluations could well track beauty judgments given that it is at least intuitively plausible that high creativity is also associated with both high stimulus complexity and high conceptual fluency. In other words, artworks, deemed to be creative, might be those that are complex yet meaningful, just as we predict (based on the PIA Model) would be the case for artworks that are deemed to be beautiful. In support of this position there is indeed some limited evidence that higher creativity judgments are associated with higher complexity ratings for consumer products (O'Quin & Besemer, 1989).

Making any firm predictions concerning the impact of stimulus complexity and conceptual fluency on creativity judgments is, however, very risky given the general absence of data and theorising on these issues. We are also aware of the need for further caution given the existence of some factor-analytic evidence from the domain of art evaluation indicating that perceived beauty and creativity may, in fact, be dissociable constructs (e.g., see Hagtvedt, Hagtvedt, & Patrick, 2008; Kozbelt, 2004). In addition, there is further evidence indicating that creativity is associated with perceived *novelty* (Diedrich, Benedek, Jauk, & Neubauer, 2015; Runco & Charles, 1993; Zhou, Wang, Song, & Wu, 2017), such that it seems plausible that for an abstract artwork to be judged as creative it would need to be surprising, lacking prototypicality and familiarity, which are all factors that are closely aligned with conceptual *disfluency*. As such, it is entirely possible that creativity and beauty judgments could show opposite associations with manipulated levels of conceptual fluency such that it is important to remain open to the possibility.

## Preliminary Norming Study

To create a suitable set of standardised materials for use in our main experiment, we first conducted a norming study that involved a large set of freely-available abstract artworks. For each of these artworks we obtained complexity and conceptual fluency ratings from a group of participants to enable us to implement effective experimental manipulations of these two factors in our main study. We report the full methodological details of this norming study below and note here that it allowed us to extract 30 artworks in total that could be separated across five different levels of conceptual fluency, with six artworks per level, three of which had been rated as “high complexity” and three of which had been rated as “low complexity”.

### Method

*Participants.* Fifty-nine adults (49 females) participated in the norming experiment ( $M_{age} = 33$  years,  $SD_{age} = 11.69$  years,  $Range_{age} = 18$  to 67 years). The participants were volunteers recruited via an opportunity sample. No exclusion criteria were imposed that restricted participation in the study. In addition, screening of the resulting data provided no grounds for post-study participant exclusion (e.g., on the basis of speeded responding or limited response variability, which might be indicative of disengaged task completion). We acknowledge the gender imbalance in the sample for this norming study. That said, we have found no evidence in the literature to suggest that gender has a significant impact on judgments of conceptual fluency or complexity for artwork. The experiment received full ethical clearance from the Psychology and Social Science Ethics Board at the University of Central Lancashire, UK.

*Materials.* A set of 150 images of abstract paintings were obtained through an internet search using “Google Images”, with these artworks covering a wide breadth of meaningfulness and complexity. The pieces of higher meaningfulness primarily depicted stylised and abstracted versions of animals, birds, landscapes and people, whereas those of

lower meaningfulness primarily involved patterns, shapes, amorphous structures and ambiguous forms. All of the selected images had been produced by lesser-known artists and were therefore deemed to be unfamiliar. This was important in order to avoid confounding effects of stimulus familiarity given the link that has been established in the literature between perceived familiarity and perceived beauty (e.g., Zajonc, 1968). We accept, however, that there is a very small chance that a participant might have previously encountered one of our chosen items, or, indeed, might have felt that an item was familiar because of its perceived similarity to a known piece of artwork. In the absence of participants' familiarity judgments for the presented artworks, we therefore need to note a degree of caution when it comes to interpreting findings.

*Design.* To obtain complexity and conceptual fluency ratings for each piece of artwork, we used a between-participants design, with 30 participants rating the artworks for conceptual fluency and 29 participants rating them for complexity. We opted for a between-participants design to avoid the possibility that rating an artwork on one dimension (e.g., conceptual fluency) might be used as a proxy for rating it on another dimension (e.g., complexity). Although potentially unlikely given the sheer number of artworks to be rated, this was nevertheless a concern since we wished to obtain uncontaminated ratings for each dimension.

*Procedure.* The experiment was conducted using “SurveyGizmo”, an online questionnaire platform. Each participant read an information sheet and responded to a question indicating their consent to proceed with the experiment. They were informed that withdrawal was not possible once they had submitted their responses, but that they could withdraw at any point prior to this by simply closing their browser. All incomplete questionnaire responses were deleted prior to analysis. The presentation of the 150 artwork pieces was independently randomised for each participant. Each artwork was presented with

a rating scale for the participant to register either a conceptual fluency judgment or a complexity judgment, dependent on the condition that they were in.

For conceptual fluency ratings participants were asked “How meaningful is this picture to you?” and for complexity ratings they were asked “How complex is this picture to you?”. They registered their responses on a sliding scale that ranged from 1 to 100. The scale end-points had descriptors (i.e., “not at all meaningful” vs. “very meaningful” for the conceptual fluency scale; “not at all complex” vs. “very complex” for the complexity scale). These end-point descriptors were reversed for approximately half of the participants in each condition to counteract any spatial biases that might impact ratings. The starting point of the slider was always set at the scale midpoint (i.e., 50).

Participants were additionally provided with the following instructions when asked for meaningfulness ratings in the conceptual fluency condition: “Meaningfulness is your own personal opinion of how meaningful a picture is to you. There is no right or wrong answer”. Likewise, when asked for complexity ratings participants were instructed as follows: “Complexity is your own personal opinion of how complex a picture is to you. There is no right or wrong answer”.

The instructions relating to meaningfulness and complexity were deliberately kept simple. We appreciate that the terms “complexity” and “meaningfulness” are open to individual interpretation, but we also note that previous research indicates that the standard custom and practice when eliciting ratings for these concepts is to leave them undefined, instead relying on people’s everyday understanding (e.g., see Lindauer & Dintruff, 1975; Snodgrass & Vanderwart, 1980; Turner & Silvia, 2006). In support of this approach Russell and Milne (1997) argue that when studying aesthetic liking it is important to avoid *over-defining* what are familiar and well-understood concepts as this can lead to spurious findings that have limited relevance to people’s natural judgments.

Participants were not given a time restriction for providing a rating for each artwork piece. However, they were informed that the rating experiment should take approximately 30 minutes to complete. Participants were debriefed at the end of the questionnaire and thanked for their time.

## **Results**

For each piece of artwork, we computed mean and standard deviation values for its rated meaningfulness and complexity, reverse scoring the ratings on those scales for which we had inverted the end-points. In all cases the final mean values were such that higher scores reflected greater meaningfulness or greater complexity.

For the purpose of the main experiment we note that the meaningfulness ratings for artworks served as a proxy for our conceptual fluency factor. We wanted to construct five linearly increasing levels of conceptual fluency, with six artworks at each consecutive level having a mean separation of around 10 points. We also wished to manipulate complexity at each level of conceptual fluency so as to have three relatively low complexity artworks that fell into the complexity range of 30 – 40, and three relatively high complexity artworks that fell into the complexity range of 50 – 60. We additionally wanted to ensure that the mean complexity of the high complexity items was broadly equivalent across all levels of conceptual fluency (and likewise for the low complexity items).

From the original 150 normed items we were able to extract 30 artworks that closely fulfilled the aforementioned criteria in order to achieve a well-controlled set of materials for use in the main experiment (see Table 1 for the mean ratings of artworks at each level of conceptual fluency and complexity). The average meaningfulness rating for the original 150 items was 46.44. As such, our middle category of conceptual fluency (40.00 – 49.99) encompassed the average score, with two categories below this middle category and two above it. We note that to obtain sufficient artworks at both the low and high complexity

levels we had to extend the desired ranges minimally (i.e., 27 – 40 rather than the planned 30 – 40, and 46 – 62 rather than the planned 50 – 60).

\*\*\*Please insert Table 1 about here\*\*\*

Table 1. Mean meaningfulness ratings and complexity ratings for artworks (with mean standard deviations; SDs) at each level of conceptual fluency and complexity.

Conceptual Fluency	Complexity	Meaningfulness Rating		Complexity Rating	
		Mean	SD	Mean	SD
< 29.99	High	27.6	22.1	58.1	20.0
	Low	25.5	20.5	33.6	20.2
30.00 – 39.99	High	36.6	25.8	57.1	18.8
	Low	36.3	21.5	37.3	20.4
40.00 – 49.99	High	45.9	25.1	60.6	19.6
	Low	45.1	27.3	35.7	19.7
50.00 – 59.99	High	54.3	27.9	61.7	22.8
	Low	53.9	25.9	34.8	17.1
> 60.00	High	67.6	24.7	52.6	25.3
	Low	62.5	23.6	34.4	19.0

*Notes:* SD values for the three artwork items in each cell of the table represent averages that are calculated from the SDs obtained from the norming study. The set of 30 abstract artworks that were selected for the main experiment are accessible at <http://dx.doi.org/10.17030/uclan.data.00000152>.

Table 1 indicates a clear mean separation across conceptual fluency levels of around 10 points on the rating scale, giving rise to five linearly increasing levels of artwork meaningfulness. We acknowledge that the mean standard deviation values at each level of conceptual fluency are quite high, which reflects the fact that there was a fair degree of individual variation in meaningfulness ratings arising for each piece of artwork during the norming study. Such individual variation in ratings at the item level is typical in studies examining people's subjective experience of artworks and other aesthetic objects (e.g., Jacobsen, 2010; Jacobsen & Höfel, 2002) and means that overall trends across several levels of a factor are often more informative than pairwise comparisons between adjacent levels of a factor, which may not reveal significant differences. In the present study the use of five increasing levels of conceptual fluency encompassing artworks with meaningfulness ratings in the mid-20s at one end of the continuum and in the mid-60s at the other end of the continuum provided a fair degree of assurance that significant trends should be detectable if conceptual fluency is causally related to aesthetic judgments.

For the complexity factor, Table 1 shows a mean separation between high and low complexity artworks of around 20 – 25 points, with the exception of the highly meaningful items, where it was only possible to achieve a mean separation in complexity of around 18 points. We suggest that this is sufficiently close to what was achieved for the other levels of conceptual fluency as to be a limited concern for the integrity of the experiment that we report below.

### **Main Experiment**

To recapitulate the aim of our main experiment, we wished to examine the impact of the complexity and conceptual fluency of abstract artworks on people's judgments of beauty and creativity. Drawing on the results of our artwork norming study, we were able to construct a highly controlled experiment that manipulated the complexity of artworks at two

levels and their conceptual fluency at five levels in order to detect any effects of these variables on aesthetic judgments.

## **Method**

*Participants.* Sixty-six adults (33 female) participated in the experiment ( $M_{age} = 34$  years,  $SD_{age} = 9$  years,  $Range_{age} = 21$  to 59 years). Participants were recruited via an opportunity sample using the online participation forum “Prolific Academic”. As with the norming study, no exclusion criteria were imposed that restricted participation in the main experiment. In addition, screening of the resulting data provided no grounds for post-experiment participant exclusion. Participants were reimbursed £1.50 for 15 minutes of participation time. G-Power was used to calculate the a priori sample size required for a 0.15 Cohen’s  $d$  (i.e., a small effect size) with 90% power (Faul, Erdfelder, Lang, & Buchner, 2007). This indicated a minimum recommended sample size of 63 participants. The experiment received full ethical clearance from the Psychology and Social Science Ethics Board at the University of Central Lancashire, UK.

*Design and Materials.* We implemented 2 x 5 fully within-participants design using the materials extracted from the norming study reported above, with key properties of the materials as depicted in Table 1. The complexity factor had two levels (low complexity: 27 – 40; high complexity: 46 – 62) whilst the conceptual fluency factor had five levels that reflected increasing gradations in the rated meaningfulness of artworks (< 29.99; 30.00 – 39.99; 40.00 – 49.99; 50.00 – 59.99; and > 60.00). Each combination of a level of conceptual fluency with a level of complexity encompassed three pieces of abstract artwork. Table 1 presents details of the descriptive data associated with each of the 10 categories of artwork. The full set of 30 abstract artworks selected for use in this experiment are accessible at <http://dx.doi.org/10.17030/uclan.data.00000152>.

*Procedure.* The experiment was run online using the SurveyGizmo platform. Each participant read an information sheet and responded to a question indicating consent to proceed with the experiment. The same withdrawal protocol was applied as in the norming study.

All participants rated each of the 30 pieces of presented artwork. There were six pieces at each level of conceptual fluency, with half of these being high in complexity and the other half being low in complexity. Participants rated each piece of artwork twice, once for beauty and once for creativity. The 30 ratings for beauty were blocked, as were the 30 ratings for creativity. The rating order for beauty and creativity was counterbalanced across participants (i.e., half of the participants conducted beauty ratings before creativity ratings and half had the opposite ordering). Furthermore, the presentation order of the 30 pieces of artwork was independently randomised by the survey program for each participant, both when they engaged in the beauty ratings and when they engaged in the creativity ratings.

As in the norming study, each participant provided beauty and creativity ratings using a sliding pointer on a 100-point scale, with the pointer commencing at the mid-point. For beauty ratings participants were asked “How beautiful is this picture to you?”, with the scale end-point descriptors (i.e., “not at all beautiful” vs. ‘very beautiful’) reversed across half of the participants to ensure counterbalancing. For creativity ratings participants were asked “How creative is this picture to you?”, with the scale end-point descriptors (i.e., “not at all creative” vs. ‘very creative’), again reversed across half of the participants. In addition, participants were given the following instructions when asked for beauty [or creativity] ratings: “Beauty [Creativity] is your own personal opinion of how beautiful [creative] a picture is to you. There is no right or wrong answer”.

Participants were not required to complete the experiment within a specific timeframe, however, they were informed that it should take approximately 15 minutes in

total. All participants were fully debriefed at the end of the experiment and thanked for their participation time.

## Results

Descriptive data relating to both beauty and creativity ratings are shown in Table 2. An alpha level of .05 was adopted for all statistical tests unless specified otherwise in the text. A rating score of between 1 and 100 was obtained for each artwork piece in response to both the beauty and creativity questions. Consistent with the norming study, we reverse scored the ratings on those scales for which we had inverted the end-points. In all cases the final mean values were such that higher scores reflected greater beauty judgments or greater creativity judgments. We analyzed beauty and creativity ratings separately using analysis of variance (ANOVA).

\*\*\*Please insert Table 2 about here\*\*\*

*Beauty.* A 2 (complexity: high vs. low) by 5 (conceptual fluency: < 29.99; 30.00 – 39.99; 40.00 – 49.99; 50.00 – 59.99; and > 60.00) repeated-measures ANOVA was conducted on the mean beauty ratings. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated for the conceptual fluency factor,  $\chi^2(9) = 61.34$ ,  $p < .001$ , and a Greenhouse-Geisser correction was therefore applied to adjust the degrees of freedom for this factor accordingly.

As predicted, the ANOVA revealed the presence of a significant main effect of conceptual fluency on beauty ratings,  $F(2.56, 166.45) = 36.45$ ,  $p < .001$ ,  $\eta_p^2 = 0.36$ . The test of within-participants contrasts indicated the existence of a highly reliable linear trend in the data,  $F(1, 65) = 53.99$ ,  $p < .001$ ,  $\eta_p^2 = 0.45$ , whereby abstract artwork was judged to be increasingly more beautiful at increasing levels of conceptual fluency. In addition, there was also evidence for the presence of a quadratic trend in the data,  $F(1, 65) = 32.36$ ,  $p < .001$ ,

$\eta_p^2 = 0.33$ , as suggested by the slight U-shaped pattern that is visible for beauty judgments in Figure 1 across levels of conceptual fluency.

\*\*\*Please insert Figure 1 about here\*\*\*

Table 2. Mean beauty ratings and creativity ratings for artworks (with standard deviations; SDs) at each level of conceptual fluency and complexity.

Conceptual Fluency	Complexity	Beauty Rating		Creativity Rating	
		Mean	SD	Mean	SD
< 29.99	High	47.2	20.0	51.7	20.2
	Low	44.1	20.3	49.8	17.4
30.00 – 39.99	High	39.6	20.1	49.5	18.4
	Low	40.6	19.8	49.1	17.6
40.00 – 49.99	High	54.9	15.0	50.5	23.0
	Low	46.5	16.5	50.2	17.9
50.00 – 59.99	High	57.9	18.9	51.2	23.6
	Low	48.6	17.3	50.3	14.2
> 60.00	High	70.2	13.8	51.5	21.0
	Low	56.7	14.8	51.3	13.4

Pairwise comparisons with a Bonferroni adjustment applied for multiple comparisons indicated a significant difference between all levels of meaningfulness (all  $p$ s < .005), with the exception of the difference between < 29.99 versus 40.00 – 49.99 ( $p = .10$ ), and the difference between 40.00 – 49.99 versus 50.00 – 59.99 ( $p = .99$ ). The failure of the latter pairwise comparison to reveal a significant difference may reflect the issue that we mentioned previously, whereby the somewhat high average standard deviations in meaningfulness ratings for items at each level of conceptual fluency may obscure the emergence of significant differences between levels. The same issue may underpin the quadratic effect arising in the test of within-participants contrasts, which may have occurred because items at the lowest level of conceptual fluency potentially had degrees of meaningfulness that overlapped with items at higher levels of conceptual fluency for some participants. Notwithstanding these caveats, we suggest that the data broadly align very well with the existence of a predicted effect of conceptual fluency on beauty ratings.

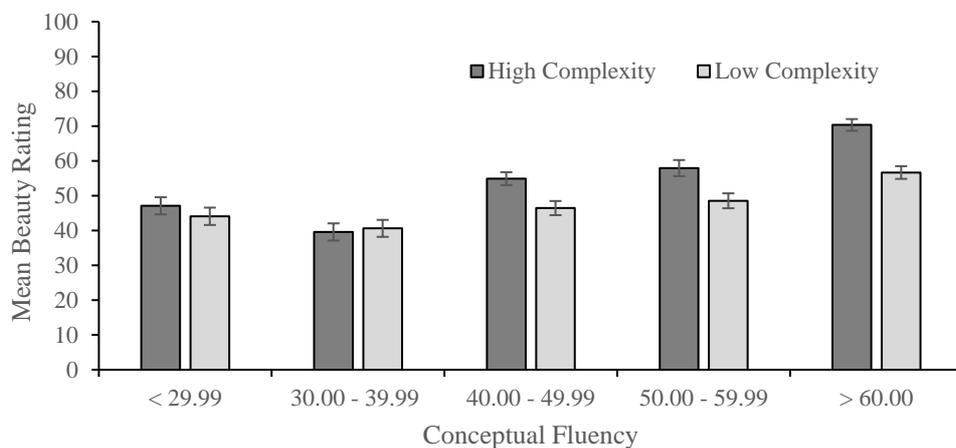


Figure 1. Mean beauty ratings for artworks as a function of conceptual fluency and complexity. Error bars represent standard errors.

Interestingly, the ANOVA also revealed the existence of a significant main effect of complexity on beauty ratings,  $F(1, 65) = 41.07$ ,  $p < .001$ ,  $\eta_p^2 = 0.39$ , with high complexity artworks ( $M = 53.96$ ,  $SE = 1.50$ ) being rated as significantly more beautiful than low complexity artworks ( $M = 47.28$ ,  $SE = 1.39$ ). Also as predicted, the effect of complexity modulated the effect of conceptual fluency, as revealed by a significant conceptual fluency by complexity interaction,  $F(4, 260) = 6.38$ ,  $p < .001$ ,  $\eta_p^2 = 0.09$ . To clarify the source of this interaction we conducted simple main effects analyses with Bonferroni corrections applied for multiple comparisons. These analyses revealed significant differences in mean beauty ratings between high versus low complexity artworks across the three highest levels of conceptual fluency, that is: 40.00 – 49.99, 50.00 – 59.99 and  $> 60.00$  (all  $ps < .001$ ). However, there was no significant difference in mean beauty ratings between high versus low complexity artworks at the two lowest levels of conceptual fluency, that is:  $< 29.99$  ( $p = .19$ ) and 30.00 – 39.99 ( $p = .68$ ). In other words, complexity appeared to impact on beauty ratings for abstract artwork pieces only at a conceptual fluency levels approximating 40 or greater. Figure 1 clearly depicts this significant interaction between conceptual fluency and complexity.

*Creativity.* A 2 (complexity: high vs. low) by 5 (conceptual fluency:  $< 29.99$ ; 30.00 – 39.99; 40.00 – 49.99; 50.00 – 59.99; and  $> 60.00$ ) repeated-measures ANOVA was conducted on the mean creativity ratings. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated for conceptual fluency;  $\chi^2(9) = 94.34$ ,  $p < .001$ , and for the conceptual fluency by complexity interaction  $\chi^2(9) = 24.45$ ,  $p = .004$ . Therefore, a Greenhouse-Geisser correction was applied to adjust the degrees of freedom in both instances.

The ANOVA revealed that there was no significant main effect of complexity on creativity ratings,  $F(1, 65) = 0.15$ ,  $p = .70$ ,  $\eta_p^2 = .002$ . Mean creativity ratings for the high

complexity pictures ( $M = 50.87$ ,  $SE = 1.94$ ) were, in fact, nearly identical to those for the low complexity pictures ( $M = 50.11$ ,  $SE = 1.30$ ), with both of these mean values being located very close to the mid-point of the creativity rating scale (see Table 2). The ANOVA also indicated that there was no significant main effect of conceptual fluency on creativity ratings,  $F(2.28, 148.07) = 0.27$ ,  $p = .793$ ,  $\eta_p^2 = .004$ , with the ratings in Table 2 again indicating that all mean creativity scores were located close to the scale mid-point. Unsurprisingly, there was also no significant conceptual fluency by complexity interaction,  $F(3.26, 212.11) = 0.11$ ,  $p = .963$ ,  $\eta_p^2 = .002$ .

### **General Discussion**

The present study set out to examine how aesthetic liking judgments are influenced by two factors that have received limited attention in the literature on empirical aesthetics, that is, conceptual fluency (i.e., the ease of extracting a meaningful interpretation from a presented stimulus) and stimulus complexity. Our research was partly motivated by our intuition that these two factors would be likely to play an important role in the aesthetic appreciation of abstract art given that such art is frequently considered to be visually challenging whilst still often being viewed positively. It therefore seemed important to manipulate these two factors simultaneously in order to unpack the way in which they might contribute to the aesthetic liking of abstract artworks.

Our research was not entirely driven by curiosity, however, but was also motivated by theoretical concerns relating to limitations that are becoming increasingly apparent with respect to the dominant theory of aesthetic liking, that is, Reber et al.'s (e.g., 2004a) processing-fluency account. This account places a key emphasis on how feelings of fluency that arise when processing a stimulus engender feelings of liking, with such processing fluency typically arising either from the visual properties of a stimulus (e.g., its symmetry, contrast or clarity; Reber, 2002; Reber et al., 2004b; Reber et al., 1998; Wurtz et al., 2008) or

from a perceiver's previous processing experience with the stimulus (e.g., prior exposure to it; Bornstein & D'Agostino, 1994). Although this processing-fluency account has been able to explain a wide range of data from research on empirical aesthetics it has been subjected to increasing challenges in recent years. There is, for example, mounting evidence that runs counter to the prediction of a monotonically increasing relationship between processing fluency and aesthetic liking (e.g., Hekkert et al., 2003; Landwehr et al., 2013). There is also growing evidence that factors that seem to be linked to *disfluency*, such as complexity and novelty, may paradoxically engender aesthetic liking (e.g., Blijlevens et al., 2012; Landwehr et al., 2011; Martindale et al., 1990).

Although attempts have been made to add auxiliary assumptions to the processing-fluency account (e.g., Reber et al., 2004) in order to reconcile the theory with anomalous findings, there is growing unease with the viability of the account such that new theories are beginning to emerge aimed at providing a more comprehensive explanation of extant data whilst also fueling new predictions. One such account is Graf and Landwehr's (2015) Pleasure Interest Model of Aesthetic Liking (PIA Model), which draws extensively on dual-process notions such as those associated with default-interventionist theories of reasoning (e.g., Evans, 2017; Evans & Stanovich, 2013). The PIA model emphasises how aesthetic liking can be driven by two processes: (i) an initial process that is automatic and stimulus-driven, delivering default aesthetic judgments akin to those captured by the traditional processing-fluency model; and (ii) a second, deliberative process that is elaborative, interpretative and perceiver-driven in nature, delivering outputs that may overturn or reinforce default preferences.

Importantly in terms to the predictions driving our research, the second process in the PIA Model is assumed often to be triggered by *disfluency* arising at the default processing stage – as might occur, for example, when a stimulus is complex. The triggering of the

deliberative stage will typically involve motivated processing aimed at deriving meaning, which will be facilitated if a stimulus is conceptually fluent despite its apparent complexity. If meaning extraction is relatively straightforward at this second stage, then this will engender enhanced aesthetic liking even in the face of apparent complexity, thereby potentially explaining the occurrence of some seemingly anomalous findings in the literature on art appreciation (e.g., Belke et al., 2006; Leder et al., 2006; Millis, 2001; Russell, 2003).

The latter processing sequence that we have outlined aligns closely with the processing sequence that we expected to occur for the complex stimuli we devised for our reported study of aesthetic liking for abstract artworks. The study manipulated stimulus complexity (high vs. low) together with the conceptual fluency of the presented artworks, which were associated with five distinct levels of predetermined conceptual fluency. Our overarching prediction was that our aesthetic liking data would reveal the existence of an interaction between stimulus conceptual fluency and complexity given that high complexity stimuli should trigger more deliberative processing than low complexity stimuli, with such deliberative processing then engendering higher liking ratings at increasing levels of conceptual fluency. Our findings revealed the existence of this expected modulation of conceptual fluency by complexity in relation to aesthetic judgments of beauty, thereby attesting to the predictive strength of Graf and Landwehr's (2015) PIA Model (see also Graf & Landwehr, 2017). Put simply, we predicted and found that people like more complex abstract artworks compared to simpler ones, but only when they can fairly readily derive meaning from these apparently complex stimuli. In cases where the extraction of meaning is more elusive, then people show reduced liking for abstract artworks and no separation in liking between complex versus simpler pieces.

These latter findings, however, were only apparent in our measures of aesthetic liking in terms of *beauty*. When it came to measures of *creativity* our findings ran counter to

predictions and simply appeared to be uniform across all factors and levels, merely residing at the mid-point of the creativity scale. This null effect is both striking and curious, especially since the complexity and conceptual fluency factors were evidently having such an impact on beauty judgments. How, then, might we explain this lack of sensitivity of our creativity measure to our manipulations, with participants seemingly defaulting to a middle-ground evaluation of “moderately creative”? In attempting a post hoc account of this finding we first note that creativity judgments are viewed as being fundamentally concerned with perceptions of *novelty* and *value* (e.g., Mayer, 1999). Since all of our presented artworks had novelty (i.e., they were selected to be unfamiliar) they presumably would all have been weighted similarly by participants on this dimension, such that novelty would have had an essentially uniform impact on creativity judgments. In terms of value, however, the only function that our artworks served was an aesthetic one relating to their perceived beauty, which means that creativity judgments might well have been expected to closely track beauty judgments. The fact that this was not the case suggests creativity judgments might, after all, have tracked novelty – and done so to a similar degree across all pieces. In support of this speculative account of our creativity data is evidence that the perception of novelty appears to be the most important determinant of creativity (Diedrich, Benedek, Jauk, & Neubauer, 2015; Runco & Charles, 1993) and that creativity and beauty also load on different factors in factor analytic studies of art appreciation (Hagtvedt, Hagtvedt, & Patrick, 2008; Kozbelt, 2004).

The lack of any effect of complexity or conceptual fluency on creativity judgments in our study is fascinating and certainly worthy of further investigation in the context of art appreciation. Notwithstanding this null effect, however, we contend that the way in which Graf and Landwehr’s (2015) PIA Model was able to predict the findings for our beauty measure is impressive, not least because the interaction prediction was both risky and arguably non-intuitive in terms of complexity driving judgments of aesthetic liking. In this

respect, we concur with Graf and Landwehr (2015), who suggest that the PIA Model is eminently applicable to understanding the perception of artworks, where higher-order cognitive processes, such as finding meaning and understanding, seem to play such an important role in art appreciation (e.g., Leder et al., 2004, 2006). As Graf and Landwehr also note, the PIA Model further supports the idea that artists may actively use disfluency as a way to disrupt any tendency by a viewer to default to the shallow processing of a piece of art (cf. Bulot & Reber, 2013).

Although our findings in relation to people's beauty judgments appear to align well with the predictions of Graf and Landwehr's (2015) PIA Model, we acknowledge some weaknesses with our study that would be worth addressing in future research in order to facilitate an even more definitive examination and replication of predicted effects. First and foremost, we reiterate the challenge of obtaining abstract artworks at distinct levels of conceptual fluency that overlap minimally in their perceived degree of meaningfulness with other levels of conceptual fluency. Individual differences in people's ability to derive meaning from artworks will typically limit what can be achieved in terms of reducing the variance in conceptual fluency judgments at different levels of this factor. One way forward might be to constrain the definition of "meaningfulness" that is provided to participants prior to eliciting judgments from them. For example, it might be useful to elicit a more *objective* judgment of meaningfulness that reflects how much it is felt that "people in general" might derive meaning from a stimulus, although in this respect we note the cautionary comments of Russell and Milne (1997), who argue that when studying aesthetic liking it is important to avoid over-defining familiar and well-understood concepts.

Second, it would be useful for future research to examine more extreme differences between stimuli in relation to the *complexity* factor. We concede that the manipulation of low versus high complexity in our study (i.e., ratings of 30 – 40 vs. 50 – 60) meant that stimuli

were closer to the midpoint of the complexity scale than might have been desirable. This issue again speaks to the challenges of developing effective materials for use in studies in the domain of empirical aesthetics. By testing an original set of 150 abstract artworks from across a wide range of styles we had hoped that our norming study would give rise to a better separation of low versus high complexity artworks for use in our main study. We were, however, also constrained by the need to identify three high-complexity and three low-complexity items at each of the five levels of conceptual fluency, which meant that there were insufficient items across each meaningfulness level to choose more extreme complexity ratings. Nevertheless, we would argue that emergence in our study of a reliable interaction between complexity and conceptual fluency for beauty judgments supports the view that our complexity manipulation was effective despite not being as extreme as we would have liked.

Third, we acknowledge that our study potentially suffered from the fact that the sample used in our norming study was not drawn from the same population as the sample in our main experiment, which derived from the Prolific Academic online survey platform. Nevertheless, no restrictions were set on participants selecting the study through Prolific Academic, such that the impact arising from the use of differing populations in our norming study and our main experiment may well have been minimized.

As a final, positive consideration that is very much in the spirit of broadening the theoretical reach of Graf and Landwehr's PIA Model of aesthetic liking, we would like to note the close parallels between the key assumptions of the PIA Model and those of the Revised Optimal Innovation Hypothesis as advanced recently by Giora et al. (2017; for former, pioneering instantiations of this hypothesis see Giora et al., 2004, 2015). The Revised Optimal Innovation Hypothesis provides another highly compelling contemporary account of the role of default processing in affecting pleasure in relation to presented stimuli. According to the hypothesis, a stimulus would be optimally innovative if: (i) it involves a *nondefault*

response to a given stimulus, which differs from the *default* response(s) associated with it, both quantitatively and qualitatively, while (ii) allowing for the automatic recoverability of the *default* response(s) related to that stimulus, so that both the default and *nondefault* responses may be weighed against each other, their similarity and differences assessable. When these aforementioned conditions prevail then the stimulus would be experienced affectively as being pleasurable.

The parallels between the Revised Optimal Innovation Hypothesis and the core assumptions of the PIA Model seem striking. Both accounts emphasise the inherent contrast between a *default* and a *nondefault* response as being crucial for the experience of aesthetic pleasure and both accounts hold that the default response needs to be available (recoverable) so that it can be considered in relation to the nondefault response so that similarities and differences can be ascertained. Indeed, Graf and Landwehr (2015) are at pains to point out that aesthetic liking judgments reflect a complex interplay between default processing and deliberative processing, with the perceiver being alert to ongoing fluctuations and contrasts in fluency/disfluency as well as the critical juxtaposition between default fluency/disfluency and fluency/disfluency arising at the deliberative stage.

Although the mapping between the PIA Model and the Revised Optimal Innovation Hypothesis may not be perfect, there seems enough common ground to suggest fruitful lines of convergence in the future. This is an exciting prospect, not least because of the already wide explanatory reach of the Optimal Innovation Hypothesis, both in its current and its former versions. It has, for example, been applied extensively to predict and explain people's understanding of figurative language (e.g., sarcasm), including the derivation of lexically coded meanings and non-coded, constructed interpretations (e.g., Giora et al., 2017; for further, related evidence see Filik, Howman, Ralph-Nearman, & Giora, 2018, this issue; Giora, Cholev, Fein, & Peleg, 2018, this issue). The Optimal Innovation Hypothesis has also

been applied successfully in predicting and explaining pleasure ratings with pictorial stimuli (e.g., Giora et al., 2004). The hypothesis can additionally capture the potential importance of cueing nondefaultness, so that nondefault meanings do not escape a comprehender's attention (for further discussion of the need to signpost nondefaultness, see Veale, 2018, this issue; see also Giora, submitted). Such cueing of nondefault meanings appears to be relatively commonplace in artwork appreciation, where an artist's use of a carefully-crafted title or subtle pointers within the artwork itself may serve to signal the presence of initially evasive, nondefault interpretations. Although our reported work did not tackle the signposting of nondefaultness, this topic is well worth examining further in the context of art appreciation. In sum, we welcome the possibility that future research might enable integration of key elements of the PIA Model and the Revised Optimal Innovation Hypothesis in order further to clarify the role of default and nondefault processing in driving judgments of aesthetic liking across a wide range of stimuli.

### **Author Note**

We thank Rachel Giora and Ruth Filik for their valuable advice during the preparation of this article. We also thank Shira Elqayam and an anonymous reviewer for their detailed and constructive feedback on an earlier version of this article. Their recommendations greatly enhanced the reporting of this research.

## References

- Ackerman, R., & Thompson, V. A. (2017a). Meta-Reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, *21*, 607–617.
- Ackerman, R., & Thompson, V. A. (2017b). Meta-Reasoning: Shedding metacognitive light on reasoning research. In L. J. Ball, & V. A. Thompson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 1-15). Abingdon, Oxon: Routledge.
- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*, 569–576.
- Armstrong, T., & Detweiler-Bedell, B. (2008). Beauty as an emotion: The exhilarating prospect of mastering a challenging world. *Review of General Psychology*, *12*, 305–329.
- Ball, L., Threadgold, E., Solowiej, A., & Marsh, J. E. (in press). Can intrinsic and extrinsic metacognitive cues shield against distraction in problem solving? *Journal of Cognition*.
- Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science*, *17*, 645–648.
- Belke, B., Leder, H., & Augustin, M. D. (2006). Mastering style: Effects of explicit style-related information, art knowledge and affective state on appreciation of abstract paintings. *Psychology Science*, *48*, 115–134.
- Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*, 279–286.
- Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.
- Berlyne, D. E. (1974). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Washington, DC: Hemisphere.

- Bertamini, M., Makin, A., & Rampone, G. (2013). Implicit association of symmetry with positive valence, high arousal and simplicity. *i-Perception*, 4, 317–327.
- Blijlevens, J., Carbon, C. C., Mugge, R., & Schoormans, J. P. L. (2012). Aesthetic appraisal of product designs: Independent effects of typicality and arousal. *British Journal of Psychology*, 103, 44–57.
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106, 265–289.
- Bornstein, R. F., & D’Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual/attributional model of the mere exposure effect. *Social Cognition*, 12, 103–128.
- Bulot, N. J., & Reber, R. (2013). The artful mind meets art history: Toward a psycho-historical framework for the science of art appreciation. *Behavioral and Brain Sciences*, 36, 123–137.
- Carbon, C. C., Faerber, S. J., Gerger, G., Forster, M., & Leder, H. (2013). Innovation is appreciated when we feel safe: On the situational dependence of the appreciation of innovation. *International Journal of Design*, 7, 43–51.
- Carbon, C. C., & Leder, H. (2005). The Repeated Evaluation Technique (RET): A method to capture dynamic effects of innovativeness and attractiveness. *Applied Cognitive Psychology*, 19, 587–601.
- Consoli, G. (2017). Commentary: Aesthetic pleasure versus aesthetic interest: The two routes to aesthetic liking. *Frontiers in Psychology*, 8, 1197.
- Diedrich, J., Benedek, M., Jauk, E., & Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, & the Arts*, 9, 35–40.

- Evans, J. St. B. T. (2017). Dual-process theories. In L. J. Ball & V. A. Thompson (Eds.), *The Routledge International Handbook of Thinking and Reasoning* (pp. 151-166). Abingdon, Oxon: Routledge.
- Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223–241.
- Faerber, S. J., Leder, H., Gerger, G., & Carbon, C. C. (2010). Priming semantic concepts affects the dynamics of aesthetic appreciation. *Acta Psychologica*, 135, 191–200.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Fechner, G. T. (1876). *Vorschule der aesthetik [Introduction to aesthetics]*. Leipzig, Germany: Breitkopf & Härtel.
- Filik, R., Howman, H., Ralph-Nearman, C., & Giora, R. (2018, this issue). The role of defaultness and personality factors in sarcasm interpretation: Evidence from eye-tracking during reading. *Metaphor and Symbol*.
- Giora, R. (submitted). How defaultness affects processing, pleasure, and cueing: The case of default constructional sarcasm and default non-constructional literalness. *Constructions & Frames*.
- Giora, R., Cholev, A., Fein, O., & Peleg, O. (2018, this issue). On the superiority of defaultness: Hemispheric perspectives of processing negative and affirmative sarcasm. *Metaphor and Symbol*.
- Giora, R., Fein, O., Kotler, N., & Shuval, N. (2015). Know hope: Metaphor, optimal innovation, and pleasure. In G. Brône, K. Feyaerts, & T. Veale (Eds.), *Cognitive linguistics meet humor research. Current trends and new developments* (pp. 129–146). Berlin, Germany: Mouton de Gruyter.

- Giora, R., Fein, O., Kronrod, A., Elnatan, I., Shuval, N., & Zur, A. (2004). Weapons of mass distraction: Optimal Innovation and pleasure ratings. *Metaphor & Symbol, 19*, 115–141.
- Giora, R., Givoni, S., Heruti, V., & Fein, O. (2017). The role of defaultness in affecting pleasure: The Optimal Innovation Hypothesis revisited. *Metaphor & Symbol, 32*, 1–18.
- Graf, L. K. M., & Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: The pleasure-interest model of aesthetic liking. *Personality & Social Psychology Review, 19*, 395–410.
- Graf, L. K. M., & Landwehr, J. R. (2017). Aesthetic pleasure versus aesthetic interest: The two routes to aesthetic liking. *Frontiers in Psychology, 8/15*.
- Hagtvedt, H., Hagtvedt, R., & Patrick, V. M. (2008). The perception and evaluation of visual art. *Empirical Studies of the Arts, 26*, 197–218.
- Hekkert, P., Snelders, D., & van Wieringen, P. C. W. (2003). “Most advanced, yet acceptable”: Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology, 94*, 111–124.
- Humphrey, D. (1997). Preferences in symmetries and asymmetries in drawings: Asymmetries between ages and sexes. *Empirical Studies of the Arts, 15*, 41–60.
- Jacobsen, T. (2010). Beauty and the brain: Culture, history and individual differences in aesthetic appreciation. *Journal of Anatomy, 216*, 184–191.
- Jacobsen, T., & Höfel, L. E. A. (2002). Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual & Motor Skills, 95*, 755–766.
- Kahneman, D. (2011) *Thinking, fast and slow*. London: Allen Lane.
- Kozbelt, A. (2004). Originality and technical skill as components of artistic quality. *Empirical Studies of the Arts, 22*, 157–170.

- Landwehr, J. R., Labroo, A. A., & Herrmann, A. (2011). Gut liking for the ordinary: Incorporating design fluency improves automobile sales forecasts. *Marketing Science*, *30*, 416–429.
- Landwehr, J. R., Wentzel, D., & Herrmann, A. (2013). Product design for the long run: Consumer responses to typical and atypical designs at different stages of exposure. *Journal of Marketing*, *77*, 92–107.
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 489–508.
- Leder, H., Carbon, C. C., & Ripsas, A. L. (2006). Entitling art: Influence of title information on understanding and appreciation of paintings. *Acta Psychologica*, *121*, 176–198.
- Lindauer, M. S., & Dintruff, D. D. (1975). Contrast effects in the response to art. *Perceptual & Motor Skills*, *40*, 155–164.
- Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *American Journal of Psychology*, *103*, 53–80.
- Mayer, R. E. (1999). Fifty years of creativity research. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 449–460). Cambridge, UK: Cambridge University Press.
- Meyer, A., Frederick, S., Burnham, T. C., Guevara Pinto, J. D., Boyer, T. W., Ball, L. J., ... & Schuldt, J. P. (2015). Disfluent fonts don't help people solve math problems. *Journal of Experimental Psychology: General*, *144*, e16.
- Millis, K. (2001). Making meaning brings pleasure: The influence of titles on aesthetic experiences. *Emotion*, *1*, 320–329.
- Muth, C., & Carbon, C. C. (2013). The Aesthetic Aha: On the pleasure of having insights into gestalt. *Acta Psychologica*, *144*, 25–30.
- Muth, C., Pepperell, R., & Carbon, C. C. (2013). Give me gestalt! Preference for cubist artworks revealing high detectability of objects. *Leonardo*, *46*, 488–489.

- Oppenheimer, D. M., & Frank, M. C. (2008). A rose in any other font would not smell as sweet: Effects of perceptual fluency on categorization. *Cognition, 106*, 1178–1194.
- O’Quin, K., & Besemer, S. (1989). The development, reliability, and validity of the revised creative product semantic scale. *Creativity Research Journal, 2*, 267–278.
- Reber, R. (2002). Reasons for the preference for symmetry. *Behavioral & Brain Sciences, 25*, 415–416.
- Reber, R., Schwarz, N., & Winkielman, P. (2004a). Processing fluency and aesthetic pleasure: Is Beauty in the perceiver's processing experience? *Personality & Social Psychology Review, 8*, 364–382.
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*, 45–48.
- Reber, R., Wurtz, P., & Zimmermann, T. D. (2004b). Exploring “fringe” consciousness: The subjective experience of perceptual fluency and its objective bases. *Consciousness & Cognition, 13*, 47–60.
- Runco, M. A., & Charles, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. *Personality & Individual Differences, 15*, 537–546.
- Russell, P. A. (2003). Effort after meaning and the hedonic value of paintings. *British Journal of Psychology, 94*, 99–110.
- Russell, P. A., & Milne, S. (1997). Meaningfulness and hedonic value of paintings: Effects of titles. *Empirical Studies of the Arts, 15*, 61–73.
- 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 & Memory, 6*, 174–215.

- Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013a). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition, 128*, 237-251.
- Thompson, V. A., Ackerman, R., Sidi, Y., Ball, L. J., Pennycook, G., & Turner, J. A. P. (2013b). The role of answer fluency and perceptual fluency in the monitoring and control of reasoning: Reply to Alter, Oppenheimer, & Epley (2013). *Cognition, 128*, 256-258.
- Topolinski, S., & Strack, F. (2009). The analysis of intuition: Processing fluency and affect in judgements of semantic coherence. *Cognition & Emotion, 23*, 1465–1503.
- Turner, S. A., & Silvia, P. J. (2006). Must interesting things be pleasant? A test of competing appraisal structures. *Emotion, 6*, 670–674.
- Veale, T. (2018, this issue). The “default” in our stars: Signposting non-defaultness in ironic discourse. *Metaphor and Symbol*.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 19*, 1235–1253.
- Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science, 17*, 799–806.
- Winkielman, P., Schwarz, N., & Belli, R. F. (1998). The role of ease of retrieval and attribution in memory judgments: Judging your memory as worse despite recalling more events. *Psychological Science, 9*, 124–126.
- Wurtz, P., Reber, R., & Zimmermann, T. D. (2008). The feeling of fluent perception: A single experience from multiple asynchronous sources. *Consciousness & Cognition, 17*, 171–184.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology, 9*, 1–27.

Zhou, J., Wang, X. M., Song, L. J., & Wu, J. (2017). Is it new? Personal and contextual influences on perceptions of novelty and creativity. *Journal of Applied Psychology, 102*, 180–202.