



OPEN ACCESS

EDITED BY

Michael B. Steinborn,
Julius Maximilian University of Würzburg,
Germany

REVIEWED BY

Haoran Xu,
Guizhou University of Finance and
Economics, China
Antonia Čačić,
University of Zagreb, Croatia

*CORRESPONDENCE

Ivana Bianchi
✉ ivana.bianchi@unimc.it

RECEIVED 17 June 2025

ACCEPTED 21 July 2025

PUBLISHED 02 September 2025

CITATION

Bianchi I, Branchini E, Uricchio T and
Bongelli R (2025) Creativity and aesthetic
evaluation of AI-generated artworks:
bridging problems and methods from
psychology to AI.
Front. Psychol. 16:1648480.
doi: 10.3389/fpsyg.2025.1648480

COPYRIGHT

© 2025 Bianchi, Branchini, Uricchio and
Bongelli. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Creativity and aesthetic evaluation of AI-generated artworks: bridging problems and methods from psychology to AI

Ivana Bianchi^{1*}, Erika Branchini², Tiberio Uricchio³ and
Ramona Bongelli⁴

¹Department of Humanities (Section Philosophy and Human Sciences), University of Macerata, Macerata, Italy, ²Department of Human Sciences, University of Verona, Verona, Italy, ³Department of Information Engineering, University of Pisa, Pisa, Italy, ⁴Department of Political Sciences, Communication and International Relations, University of Macerata, Macerata, Italy

This paper contributes to the debate on creativity, art, and artificial intelligence (AI) by integrating insights from cognitive psychology and empirical aesthetics into the field of AI, with the goal of inspiring novel empirical research. We focus on two main topics. First, we examine the indices used in psychology to operationalize *creativity* in closed-ended and open-ended tasks, with the aim not only of demonstrating the multidimensionality involved in defining creativity, but also of stimulating reflection on the benefits that might arise from developing a similar standard set of indices to test AI scoring models for assessing creativity (of both human and AI-generated responses). Second, we focus on the situation in which the creative products generated by AI are works of art, and on their aesthetic evaluation by non-expert human observers. Bridging the literature developed in psychology of art and empirical aesthetics with the literature on AI, a number of questions emerge, regarding the bias about the “expected style” of AI-generated art, and possible variables that play a role in aversion to AI-generated art. They all suggest possible future empirical research directions.

KEYWORDS

human creativity, AI creativity, human and AI-generated artworks, AI-generated art aversion, pleasure and interest model of aesthetic appreciation

1 Introduction

There is considerable debate in current journals on creativity and artificial intelligence (AI)—the joint search of “creativity” and “AI” in Google scholar returned 18,200 scientific articles in the period 2020–2025 (data accessed on May 27, 2025). The discussion is multifaceted and touches on various issues (for an overview, see [Elgammal, 2019](#); [Cetinic and She, 2022](#); [Du Sautoy, 2020](#); [Liu et al., 2025](#); [Manovich, 2022](#); [Schröter, 2019](#); [Yang and Xu, 2025](#)). As [Arielli and Manovich \(2022\)](#) point out, many of these questions stem from an anthropocentric perspective on creativity that assumes that AI mimics human performance whereas, conversely, conceptualizing creativity in relation to AI within a non-human paradigm might illuminate the debate in new ways (see also [Landers, 2025](#); [Mazzone and Elgammal, 2019](#)).

We see the advantage of exploring the topic from a non-anthropocentric perspective. At the same time, however, we believe that there are still ample opportunities to link research developed in psychology on human creativity to the ongoing debate on creativity and AI, in order to stimulate new perspectives from which to approach the topic and new empirical

research. This paper contributes to these goals by focusing on two topics: first, the definition and operationalization of “creativity” in the psychology and AI literature; and second, the aesthetic appreciation of AI-generated art. In this context, aesthetic appreciation is defined as the response of liking (in terms of beauty and other possible scales, such as complexity, realism, effort, authenticity, etc.) in ordinary observers, as is commonly defined in the literature on the psychology of art. In other words, we are not dealing with evaluations made by artists, art critics, or other professionals of this sort. We are simply interested in comparing the reactions of naïve observers to human- and AI-generated artistic products, which is also the goal of all the literature cited in this paper. Understanding the cognitive biases behind the public’s common reaction to AI-generated artistic works is a timely topic that allows us to discover interesting aspects of the human mind and identify areas for improvement from the perspective of AI developers.

In reviewing the literature on these two topics, we adopt a specific point of view and do not address other topics, such as the definition of art and works of art. These are interesting subjects, but totally beyond the scope of this paper. Our overarching goal is to apply ideas and methods from cognitive psychology and empirical aesthetics to the current debate about AI creativity and AI-generated images that mimic artistic works. We also aim to identify new questions and future directions for empirical studies inspired by this perspective.

Before developing these two main topics in detail (see sections “Adopting indices used to assess human creativity in psychology and cognitive science to develop metrics/models for assessing AI-generated creativity” and “Empirical Aesthetics”) some preliminary considerations are necessary.

2 Some preliminary clarifications

2.1 Product versus process

First, it is crucial to clarify that the perspective adopted in this paper in no way implies that an AI capable of learning to produce highly regarded new paintings in the style of Mondrian (Noll, 1967), to generate Chinese landscapes so convincingly as to deceive observers (Xue, 2021), or to create a beautiful “next Rembrandt” (Sovhyra, 2021) follows the same *processes* as human artists. Even AICAN (Artificial Intelligence Creative adversarial Network), that is, an AI algorithm designed not simply to emulate an established style (as GAN algorithms tend to do), but to autonomously generate something new based on existing works (Elgammal, 2019; Elgammal et al., 2017), does not intend to reproduce the *process* that leads human artists to create their artworks.

Any human-generated artwork is the result not only of the human cognitive apparatus that supports its creation, but also of the personal and cultural (i.e., historical, conceptual, and symbolic) journey that led a particular artist to create a particular artwork. To put it in Yang and Xu (2025, p. 2) terms, human creativity is typically a highly personalized process, often emotion-driven and full of uncertainty and uniqueness (García, 2024), with inspiration often being spontaneous and unpredictable. Consistent with this, it often has a self-expressing meaning and is not necessarily oriented towards external judgment (i.e., has a social meaning), while AI creativity

relies on a computation-driven systematic process of analyses and reorganization of vast amounts of data (see also Runco, 2025). AI simply mimics the final *product* on a formal level (see also Arielli and Manovich, 2022; Hertzmann, 2020), corresponding to the minimization of a cost function of similarity to a known distribution of images (Goodfellow et al., 2020; Rombach et al., 2022) or, in the case of ICAN, generates a novel object that represents an “optimal” point between imitation and deviation from existing styles. AI has been successfully applied to design new paintings (Noll, 1967; Sovhyra, 2021; Xue, 2021), architecture with precise styles (e.g., Newton, 2019), to write poetry (e.g., Köbis and Mossink, 2021), novels (e.g., Green, 2020) and to complete the unfinished musical works of masters (e.g., Beethoven’s unfinished 10th Symphony was completed in “Beethoven X—The AI Project,” a collaboration between musicologists, composers, and computer scientists who fed the AI all of Beethoven’s existing works and sketches for the 10th Symphony; similarly, Schubert’s Unfinished No. 8 Symphony was completed by Huawei’s AI, which worked with Lucas Cantor to arrange these melodies into an orchestral score in Schubert’s style). In general, the creation of art using AI is nowadays a common practice for many video artists and sound artists, and even in film productions.

Since the focus in our paper is on the product, not on the process, questions that are interesting from a process point of view—for example, concerning consciousness and intentionality (Aru, 2025; Manu, 2024; Mikalonyté and Kneer, 2022; Moura, 2024; Redaelli, 2025)—are not relevant to the analyses addressed here.

2.2 The role of previously learnt knowledge

The second clarification concerns the role of previously learned knowledge in AI creativity. All the convincing products cited above come from the recombination of features, or configurations of features, learned by the AI during an initial training phase. This training phase, carried out on large databases, allows the AI to learn deeply, leading to new instances (i.e., not present in the original database) of the same types of objects used in the training.

The emphasis on the role of prior knowledge in AI-generated artefacts is often explicitly contrasted with the “novelty” that characterizes human creative responses and is used to deny the creative nature of the AI-generated products (e.g., Runco, 2023, 2025). However, framing the analysis in these terms overlooks the importance of prior knowledge in various theoretical models of creativity (e.g., Chaudhuri et al., 2024; Meyer and Pollard, 2006; Tromp and Glăveanu, 2023; van Welzen et al., 2024) and also in specific training procedures used to stimulate creativity (e.g., Birdi, 2016, 2020; Huo, 2020; Jones-Chick et al., 2022; Kienitz et al., 2014; Ritter et al., 2020; Sun et al., 2016). These procedures encourage individuals to make unusual associations between information “stored” in their knowledge bag. Unusual associations are typically thought of as remote associations, that is associations between elements that are “remote” in our usual representational network (for the importance of remote associations in the production of human creative responses, see Mednick, 1962; Kenett, 2019; Kenett and Faust, 2019; Olson et al., 2021; Toivainen et al., 2019; Wu C. L. et al., 2020). Encouraging the search for remote (rather than close and ordinary) associations is usually achieved by changing the default way in which individuals access their knowledge,

for example, by asking them to focus on the properties of the parts of the object (or problem or situation) rather than on the properties of the object as a whole (Huo, 2020), or to turn the object's properties in their opposites (e.g., Bianchi and Branchini, 2023; Branchini et al., 2021). However, in each case these strategies operate on the knowledge of the thinker, that is, on previously learnt material. Also atypical and unusual (because distant) connections presuppose linking concepts, features, and elements that were already present in one's mental organization. In other words, just as AI relies on its training data, humans draw on their accumulated knowledge, experience (including training experience) and cultural heritage as input data when creating something new. By emphasizing the importance of prior knowledge in human creativity, we are not denying the existence of differences in the processes that lead to the recombination of this prior data. As we clarified in the previous section, we give for granted that these processes differ. We simply want to acknowledge the importance of prior knowledge as a fundamental ingredient in both human- and AI-generated creative products. Once this has been clarified, then a new question raises, that is, the question of how, if at all, the idea of transcending the accumulation of existing knowledge, matter, and concepts applies to products made by AI. This idea applies to human art, which transcends existing knowledge by offering unique perspectives, challenging established ideas, and prompting new ways of seeing the world, ourselves, and art itself. Can this apply to AI-generated art? How can we identify and measure this component of transcendence?

3 Adopt indices used to assess human creativity in psychology and cognitive science to develop metrics/models for assessing AI-generated creativity

Many classic definitions of creativity developed within cognitive science state that creativity is the ability to generate ideas that are both novel/original and effective, that is, appropriate and useful (e.g., Amabile and Tighe, 1993; Beghetto and Kaufman, 2022; Bruner, 1962; Kaufman and Sternberg, 2010; Mumford, 2003; Runco, 2012; Runco and Jaeger, 2012; Simon, 2001; Stein, 1953; Sternberg and Lubart, 1999). In line with this definition, psychologists have modelled the creative process as an interaction between a generative component—which supports the production of ideas—and an evaluative component—which concerns the selection of the ideas that seem most promising in terms of feasibility and potential success (e.g., Basadur et al., 1982; Ellamil et al., 2012; Finke et al., 1992; Hao et al., 2016; Rietzschel et al., 2019; Zamani et al., 2023).

Interest in the potential of AI to support the *generative phase* is certainly a major theme in the literature on AI, widely focused on developing and testing AI creative performance in comparison to human creative performance (e.g., Hubert et al., 2024; Ismayilzada et al., 2025; Koivisto and Grassini, 2023; Stevenson et al., 2023; Tian et al., 2023).

However, another interesting question is whether AI can also support the *evaluative phase*, or whether humans remain the gatekeepers, assessing the value and potential of AI creations. The latter position is suggested, for example, by Agrawal et al. (2017), Karimi et al. (2018), Magni et al. (2024), van Esch et al. (2019) and von Krogh (2018). At the same time, promising tests have been carried out

in recent years on the possibility of using AI to assess the creativity of human-generated responses, in both verbal (e.g., Acar et al., 2021; Beaty and Johnson, 2021; Buczak et al., 2023; Dumas et al., 2021; Luchini et al., 2025; Patterson et al., 2024; Stevenson et al., 2020, 2023) and visual (e.g., Acar et al., 2025; Cropley and Marrone, 2025; Grassini and Koivisto, 2025; Patterson et al., 2024) creativity tasks. In some recent works the automated assessment was extended, comparatively, also to AI-generated responses (e.g., Chakrabarty et al., 2025; Hubert et al., 2024; Ismayilzada et al., 2025; Kern et al., 2024; Orwig et al., 2024; Stevenson et al., 2023). The goal in all cases is to develop automated scoring systems that mimic human classification of responses.¹

In the next section, rather than comparing the performance of humans and AI on specific tasks, we will take a step back and consider the *types of tasks* and *indices* that are typically used in psychology to assess human convergent and divergent creativity (see Table 1). We will then discuss how this can inspire the development of new automated metrics and models to evaluate creativity in AI- and human-generated responses.

3.1 Transferring psychological constructs from convergent and divergent thinking tests to AI-generated responses

Classifying responses as creative or not (or as creative to varying degrees) has been a challenging topic for psychological research since it is not straightforward. The complexity of the classification becomes particularly evident when *open-ended tasks* are involved, that is, tasks that admit many alternative solutions and that relate to divergent thinking, while it is easier when creative convergent thinking and *closed-ended tasks* are concerned (for these definitions see Guilford, 1967a).

Closed-ended tasks are characterized by having one correct solution which is usually not obvious and not easily accessed due to some unnecessary constraints introduced by the human mind in the

1 For example, Cropley and Marrone (2025) used a Convolutional Neural Network (deep learning) to automate the scoring of figural creativity. They focused on the Creative Thinking-Drawing Production test (TCT-DP; Urban and Jellen, 1996; Urban, 2005), in which participants are presented with squared frames containing incomplete and irregular figural fragments (e.g., a dot, a dashed segment, a circle) and are asked to complete the drawings. Based on the TCT-DP manual, 14 criteria must be used to score responses, such as connection by theme (do participants integrate the individual elements into a thematic whole?), the use of perspective, the presence of boundary-breaking (do participants draw elements that extend beyond the squared frame?), or unconventional use of the figural fragments/elements. Cropley and Marrone (2025) trained AI using a sample of responses produced by participants and scored by two independent raters (with inter-rater agreement always >0.90). They tested various AI models of scoring, which varied from two levels of creativity (low and high) to seven levels of creativity (corresponding to the seven classes defined in the TCT-DP test manual). Despite the small dataset of images used by AI in the training phase ($N = 414$), very good accuracy scores for the automated scoring emerged, along with high levels of inter-rater agreement with human raters.

TABLE 1 Summary of the measures typically used in psychology to assess human creativity in relation to closed- and open-ended tasks.

Measure	Closed-ended tasks (CET)	Open-ended tasks (OET)
Success rate	✓	–
Fluency (Number of attempts made, in CET/of ideas produced, in OET)	✓	✓
Metacognitive feelings of certainty/uncertainty (feeling of confidence or rightness)	✓	✓
Feeling of being close to the solution (i.e., feeling-of-warmth) in CET/close to find a good idea in OET	✓	✓
Analyses of drawings produced during the search phase	✓	✓
Neurophysiological states accompanying or preceding the emergence of the solution in CET/accompanying the idea generation in OET	✓	✓
Eye-tracking movements	✓	✓
Flexibility (number of categorically different ideas)	–	✓
Originality:		
- Quantitatively (number or proportion of unique responses)	–	✓
- Qualitatively (expert ratings using the Consensual Assessment Technique [CAT])	–	✓
Usefulness or utility (real-world applicability of an idea):		
- Feasibility, i.e., the practicality of an idea (expert ratings using the CAT)	–	✓
- Value, i.e., the effectiveness of an idea in achieving a goal or reducing economic costs (expert ratings using the CAT)	–	✓

See the main text of the paper for a discussion of these two types of tasks and methods.

initial representation of the situation/problem. Typical closed-ended tasks used in psychology are visual–spatial problem-solving tasks (e.g., Gilhooly and Murphy, 2005; Webb et al., 2021) or verbal problem-solving tasks (Webb et al., 2021) (Figure 1).

With closed tasks, there are simple and accessible ways to assess participants’ performance, the most straightforward being the success rate, that is, the ability of the participant to find the correct solution. These kinds of problems are structured to push the observer towards an initial representation that cannot lead to the correct solution. Finding the correct solution manifests that the problem solver has been able to overcome the initial fixation and restructure their representation of the problem in a novel and less default—in this sense more creative—representation (Danek, 2018; Danek et al., 2020; Graf et al., 2023). Various other measures of creativity can complement the success rate, such as the number of attempts made, for example. In convergent creativity and closed-ended tasks, the best performance is associated with reaching the solution as soon as possible, that is, the less attempts made, the more creative one is considered (e.g., Bianchi et al., 2020; Branchini et al., 2016; Öllinger et al., 2017). Other additional indexed aim to reveal the mental path followed by the participant while searching for the solution. For example, by analyzing the linguistic expressions used when thinking aloud in small groups, psychologists can track metacognitive aspects such as participants’ feelings of certainty or uncertainty, and how close they feel to finding the solution (Danek and Salvi, 2020; Danek and Wiley, 2017; Laukkonen et al., 2021; Salvi et al., 2016; Threadgold et al., 2018; Webb et al., 2016; Zedelius and Schooler, 2015). Other indexes concern the features of the drawings participants sketch while searching for a solution (e.g., Bianchi et al., 2020; Branchini et al., 2016; Fedor et al., 2015; Öllinger et al., 2017), or neurophysiological states that co-occur with or precede the emergence of the solution (e.g., Danek et al., 2015;

Danek and Flanagin, 2019; Salvi and Bowden, 2024), and eye movements (e.g., Bilalić et al., 2021; Ellis et al., 2011; Ellis and Reingold, 2014; Knoblich et al., 2001; Xing et al., 2018, 2019). Collecting all this information requires technical mastery and can be time-consuming, but it is methodologically sound in terms of objectivity. All these measurements (except eye tracking and neurophysiologic responses) could be in theory applicable in the context of analyses of AI creative performance. For example, we could present puzzles to Large Language Models / Vision Language Models that they could solve using chain-of-thought and explanations, and in this way could track the evolution of their attempts till they reach the solution.

When responses to an open task are concerned, the situation becomes more complicated also for the analyses of human creativity. Classic examples of open-ended tasks used in psychology studies are the Alternate Uses Task (AUT), or the Five Sentence Creative Story task (Figure 2), where 2–3 min time limits are usually set. The creation of an artwork is also an obvious example of open-ended task, since there are many alternative possible solutions.

Assessment of creativity measurement with this kind of tasks is more complicated, since there is no correct response and therefore success rate does not represent a meaningful measure of performance. The same measures used to capture participants’ mental paths with closed-ended tasks can be also used for open-ended tasks (e.g., analyses of drawings produced in the search phase, analyses of thinking aloud processes, metacognitive feelings—see Ball and Christensen, 2009, 2020; Ball et al., 2021; Christensen and Friis-Olivarius, 2020), but with regard to the classification of the product, the following measures are usually considered: *fluency*, *flexibility*, *originality* and *utility* (see Beaty and

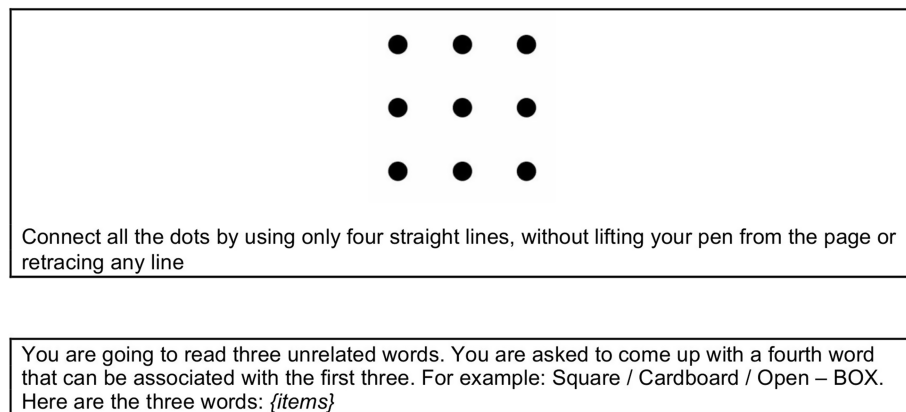


FIGURE 1

Two classic closed-ended tasks used in psychology: a visuo-spatial task (top box), i.e., the nine dots problem (Maier, 1930), and a verbal task (bottom box), i.e., the remote associates test (Mednick and Mednick, 1967).

List as many creative uses of {items} you can think of. Try to be creative as there is no right or wrong answer.

You will be given three words (e.g., car, wheel, drive) and then asked to write a creative short story that contains these three words. The idea is that instead of writing a standard story such as "I went for a drive in my car with my hands on the steering wheel.", you come up with a novel and unique story that uses the required words in unconventional ways or settings. Write a creative short story using a maximum of five sentences. The story must include the following three words: {items}.

FIGURE 2

Two classic open-ended tasks used in psychology: the alternate uses task (Guilford, 1967b) (top box) and the five sentence creative story task (Prabhakaran et al., 2014) (bottom box).

Johnson, 2021; Bellaiche et al., 2023a; Brosowsky et al., 2024; Forthmann et al., 2020; Long and Pang, 2015; Primi, 2014; Primi et al., 2019; Reiter-Palmon et al., 2019; Silvia et al., 2008; Snyder et al., 2019; Weiss and Wilhelm, 2022).

Fluency refers to the number of ideas produced by the participant. Although fluency alone does not define creativity (e.g., Rietzschel et al., 2006), the assumption is that the more ideas produced, the higher the possibility that a creative idea emerges, especially because the first ideas are the less creative and the more the person persists in ideation, the more they arrive at more creative ideas (Lucas and Nordgren, 2020).

Flexibility refers to the number of categorically different ideas. The assumption is that the more different categories involved, the more creative the person is. For instance, in the Creativity Assessment Packet (Williams, 1993), where a series of drawings must be produced starting from fragments of lines or shapes, flexibility is measured in terms of different categories of objects drawn. If a person draws a living being in one drawing, a landscape in another, a symbol in yet another, and a useful object in another, this is considered an indication of flexibility. A person whose drawings fall within the same category type (e.g., landscapes)

demonstrates less flexibility. Likewise, in AUT, response patterns classified as more flexible are those that traverse various conceptual spaces—e.g., using a “brick” as a paperweight, step to gain height, and sidewalk chalk, rather than imagining uses which variously concern holding things in place (e.g., a paperweight, bookend, doorstep) (Hass, 2017; Nijstad et al., 2010).

Originality refers to the degree of uniqueness of the idea produced, and the evaluation of this aspect is more challenging. Originality can be described quantitatively, in terms of the numbers of unique responses (initially formalized by Wallach and Kogan, 1965). The assumption is that the smaller the proportion of participants who come up with the same idea, the more original the idea is (e.g., DeYoung et al., 2008; Hocevar, 1979; Putman and Paulus, 2009; Webb et al., 2021—for a critical view of uniqueness scoring and ways to address its problems see Silvia et al., 2008). However, originality is traditionally measured qualitatively using Amabile's (1983, 1996) Consensual Assessment Technique (CAT). In this technique, domain experts independently rate responses on a scale ranging from “not at all creative” to “very creative” (Silvia et al., 2008; see also Baer and Kaufman, 2019; Kaufman et al., 2008).

Usefulness, or *utility*, corresponds to the real-world applicability of an idea and is defined in terms of both *feasibility*—that is, the practicality of an idea (Poetz and Schreier, 2012; Rietzschel et al., 2010, 2019)—and *value* (Litchfield et al., 2015)—which lacks a universal definition but is mostly related to the *effectiveness* of an idea in achieving a goal or in terms of *economic costs* (e.g., Ford and Gioia, 2000). Usefulness is also usually measured with the CAT. However, because usefulness is difficult to score, raters are generally only asked to score originality, despite usefulness playing an important and distinct role in evaluating creativity (Diedrich et al., 2015; Rietzschel et al., 2019).

The problem with CAT scoring is, first, its reliability. While some studies have proven its validity (Amabile, 1982; Kaufman et al., 2007; Long and Wang, 2022; Myszkowski and Storme, 2019), others have highlighted the impact of personal or cultural biases on this subjective classification (e.g., Blair and Mumford, 2007; Dailey and Mumford, 2006; Ivancovsky et al., 2019; Toh and Miller, 2016).² Additionally, the consensus reached depends on the raters involved (e.g., Cseh and Jeffries, 2019; Forthmann et al., 2017; Plucker et al., 2019; Reiter-Palmon et al., 2019; Zhou et al., 2017). Second, the CAT technique certainly requires time and efforts on the part of the raters, and costs for the researches. For this reason, as mentioned at the beginning of this section, automated scoring methods have recently been developed that have the potential to assist researchers and/or educators in assessing human responses. The classic procedure is to train the AI using a sample of human-generated responses scored by independent raters, and then to test different machine learning models for automated scoring.

Many of these studies have shown that AI scoring is as accurate and reliable as human raters (e.g., Beaty and Johnson, 2021; Cropley and Marrone, 2025; Dumas et al., 2021; Kern et al., 2024; Hubert et al., 2024; Luchini et al., 2025; Sun et al., 2024). However, when we compare these studies, a question emerges: What criteria are used to measure creativity? Stevenson et al. (2023) compared human and open AI performance on the GPT-3 and GPT-4 in the Alternate Use Test (AUT), evaluating responses in terms of *originality*, *usefulness*, *surprise*, and *flexibility*. The same task was used by Hubert et al. (2024) in addition to two other open-ended tasks (the Consequences Task and the Divergent Association Task)

to compare human and GPT-4 performance in terms of *fluency*, *originality* and *elaboration*. Ismayilzada et al. (2025) used the Five Sentence Creative Story task to compare the performance of humans and of 60 different Large Language Models in terms of *novelty*, *surprise*, *value*, *lexical* and *semantic diversity*. The Five Sentence Creative Story task was also used by Orwig et al. (2024) to compare the performance of humans and GPT-3 and GPT-4, but the responses were scored in terms of a general *creativity* rating. Guzik et al. (2023) compared ChatGPT (GPT-4) and human creative performance at the Torrance Tests of Creative Thinking by analyzing *fluency*, *flexibility* and *originality*.

A standardized system of creativity indices (dependent variables) would allow researchers to more easily compare the results of their studies and effectively compare the generative performance of different AI models with respect to different tasks. Furthermore, since some cited papers use AI to generate and score responses (Hubert et al., 2024; Ismayilzada et al., 2025; Kern et al., 2024; Orwig et al., 2024), reaching a consensus on the definitions used to operationalize fluency, flexibility, originality/novelty, feasibility, and value would help identify the parameters used in these automated AI scoring systems. We are not suggesting that this is an easy goal to achieve—on the contrary—only that it could be valuable. The brief overview provided in this section will hopefully also help to highlight the risks of using a generic “provide a rating of creativity” prompt (e.g., Orwig et al., 2024; Seli et al., 2025), since the definition of creativity is multifaceted and encompasses many dimensions—as reflected in the struggle of cognitive scientists to find appropriate ways to measure it. Using precise prompts when defining the scoring criterion seems preferable, both for human raters and AI models. Several studies have gone in this direction (e.g., Hubert et al., 2024; Ismayilzada et al., 2025; Kern et al., 2024; Stevenson et al., 2023).

4 From empirical aesthetics to AI

Let us now specifically focus on the situation in which the “creative” products generated using AI are artworks and on their aesthetic evaluation by human observers. Relatively little is known about (a) whether people can accurately attribute what they observe to the category of human-generated art versus AI-generated art (e.g., Chamberlain et al., 2018; Gangadharbatla, 2022; Hong and Curran, 2019; Samo and Highhouse, 2023; Velásquez-Salamanca et al., 2025) and (b) how people aesthetically evaluate AI-generated artworks (e.g., Agudo et al., 2022; Bellaiche et al., 2023b; Chiarella et al., 2022; Gangadharbatla, 2022; Hong and Curran, 2019; Horton et al., 2023; Messingschlager and Appel, 2025; Neef et al., 2025; Ragot et al., 2020; Velásquez-Salamanca et al., 2025; Wu Y. et al., 2020). However, it is not difficult to see how these are becoming hot topics for research. In a recent empirical study, Yang and Xu (2025) explored the core dimensions of AI creativity *from the audience's viewpoint*, uncovering audience biases against AI creativity. Namely the more AI was said to be involved in a creative process, in collaboration with a human being, the less the results were perceived as deep, authentic and attractive. However, it was perceived as more original. Here, we will add to these biases about AI creativity, other “biases” that emerge from a broad review of existing literature. For each of them, we will highlight open questions that may inspire interesting directions for future research aiming to discover the cognitive sources of these biases.

² Familiarity with the topic, for instance, can lead to overestimating the positive benefits and outcomes of an idea while underestimating the resources needed to implement it (e.g., Dailey and Mumford, 2006). Experts demonstrate higher interrater agreement in their evaluations than novices (e.g., Kaufman et al., 2008), but they also have a better understanding of the benefits of continuous (incremental) innovative ideas and more resistance toward discontinuous (radical) innovations, which are often better appreciated by non-experts (Moreau et al., 2001). The evaluative outcome also depends on personality factors. For instance, individual risk aversion and ambiguity aversion (Blair and Mumford, 2007; Toh and Miller, 2016) are associated with the selection of less creative ideas, while openness to experience (Silvia, 2008) correlates with the selection of more creative ideas. Information about the identity of the producer can also bias evaluators' assessments: evaluators often rely on perceptions of the producer's age, mental and emotional stability, and “genius” to guide their creativity evaluations (e.g., Baas et al., 2016; Rietzschel et al., 2019; Simonton, 2004, 2014). Judgments also depend on evaluators' emotional states, whether positive, negative, or neutral (e.g., Mastroia et al., 2019).

4.1 The “typical style” of visual artworks attributed to AI

In various studies in which participants were exposed to images of both abstract and figurative paintings purportedly generated by either AI or humans, it was found that figurative images were more often attributed to humans, whereas abstract images were attributed to AI (e.g., Gangadharbatla, 2022, study 1). Alternatively, some findings suggest that figurative art was attributed indifferently to AI and humans, while abstract art was preferentially attributed to AI (see Chamberlain et al., 2018, study 1).

Empirical aesthetic research has shown that non-art experts tend to have a clear preference for figurative art over modern art, which is more often appreciated by individuals with greater familiarity with and expertise in art (Bimler et al., 2019; Knapp and Wulff, 1963; Leder et al., 2012; Leder and Nadal, 2014; Mastandrea et al., 2011, 2021; Pihko et al., 2011). The bias toward attributing abstract art to AI suggests an implicit appreciation of human art (since figurative art is generally preferred) and a devaluation of AI art (since abstract art is less preferred). This is further supported by studies in which aesthetic judgments of AI-generated compared to human-generated art were explicitly solicited, as discussed in the following section. However, the aspect we want to highlight here is that these attributions suggest *that people implicitly assume that abstract art is the typical “style” of AI-generated images*. This conclusion is consistent with some findings emerged from Chamberlain et al. (2018), where participants were first asked to attribute images of artworks to either humans or AI and later to explain “How did you decide if a work was computer-generated?” (free-response task). Thematic analyses of participants’ responses revealed that there were relatively few references to content or intentionality; instead, responses focused primarily on surface and structural aspects of the artworks. Namely, the most common justifications for categorizing images as computer-generated were bright or artificial colors and rigid, straight, regular shapes and lines. Conversely, human-generated images were classified as such based on their appearance of being handmade, as evidenced by irregular, imperfect lines and a larger distribution of orientations, as well as the presence of brushstrokes (for this last aspect, see also Fuchs et al., 2015). Similarly, Noll (1967) and Schröter (2019) found that ordered and regular distributions of elements were associated with AI-generated art.

Gangadharbatla (2022, p. 15) suggests that this typical “style” of AI (associated more with simple and regular composition) may have originated from early experiences of AI generation, with algorithms that were not as sophisticated as current ones, leading people to expect that AI could only produce simpler (i.e., abstract) artworks, not compelling figurative artworks. However, following projects such as the next Rembrandt (Sovhyra, 2021) and the development of databases of high-resolution human faces and landscapes produced by the latest generation of AI algorithms (Rombach et al., 2022), this limitation has been overcome. It seems promising to investigate whether participants exposed to these updated types of AI-generated images—either because of their professional or educational background, or because they are exposed to images such as the next Rembrandt during a training phase—still exhibit the *abstract = AI versus figurative = human* equation, or whether this association would be significantly reduced. Such data would help shed light on the roots of the prototypical idea of AI-generated art

that people seem to hold. If it is simply due to a mere exposure effect, this should be a very transient bias that disappears after exposure to a training session that allows participants to acknowledge the possibility of AI generating realistic figurative images. If the bias persists, it would suggest that the bias is related to other factors (perhaps linked to a broader technology vs. anthropological cognitive dimension) that would be interesting to understand.

4.2 The diminished value of AI-generated art when compared to human-generated art

Psychology studies on the aesthetic evaluation of human-generated artworks have shown that people’s evaluation can be modified by providing contextual information, such as historical information about the artist (e.g., Fischinger et al., 2020), the style of the artwork (e.g., Russell, 2003), the intentional message the artist wanted to convey (e.g., Millis, 2001), the thought process involved in the innovation displayed by the artwork (Bianchi et al., 2025), and the skills required to create the artwork (e.g., Hodge, 2012). This information increases non-art experts’ understanding of and interest in the artwork, more than the pleasantness of the artwork (e.g., Belke et al., 2006; Bianchi et al., 2025; Jucker et al., 2014, experiments 2a, 2b; Leder et al., 2006, experiment 1).

Research has only just begun to connect the fields of empirical aesthetics and AI, but the role of contextual factors also seems central for research on AI-generated art. Studies comparing evaluations of human-generated and AI-generated art have found that providing contextual information about the creator (human or AI) changes participants’ appreciation of the artwork. This effect is typically assessed using experimental designs in which the same images are presented in two different conditions: one in which participants are unaware of the authorship and another in which they are informed that the artwork was created by a human or by “technical” tools; or by presenting the same images in one condition labelled as AI-created, and in another condition as human-created. These studies show, for example, that viewers rated images labelled as having been created in Photoshop as less aesthetically pleasing than the same images labelled as having been taken from an art gallery (Kirk et al., 2009). Similarly, participants rated the same images of painting as more beautiful when they were attributed to human creators than when they were said to have been generated by robots (Chamberlain et al., 2018; Di Dio et al., 2023) or AI machine learning algorithms (e.g., Hong and Curran, 2019; Horton et al., 2023; Millet et al., 2023; Ragot et al., 2020; Wu Y. et al., 2020). The bias against AI-generated art persists even when it is emphasized that the artwork was created in a collaborative production between humans and AI (Horton et al., 2023, study 6; Messer, 2024; Yang and Xu, 2025) and is not exclusive to the visual arts. Participants rate the quality of a music lower when they were informed it was AI-generated rather than human-made (e.g., Agudo et al., 2022; Hong et al., 2020, 2022; Moffat and Kelly, 2006; Shank et al., 2023). Similar results were observed in dance choreography (Darda and Cross, 2023). Neef et al. (2025) suggest that lower appreciation of AI-generated art emerges in particular when it is compared to human-generated images and not independently of this comparative framework (this is also highlighted in Horton et al., 2023).

In the following subsections we will focus on visual art, we outline five possible factors underlying this aversion that have been identified in

various literature (leaving aside the role of individual differences) and that prefigure possible promising lines of future research to be developed.

4.2.1 Is the aversion mitigated by the artistic style of the painting considered?

Aversion to AI-generated art may stem from a broader phenomenon known as “algorithm aversion” in decision making (Castelo et al., 2019; Dietvorst et al., 2015; Shaffer et al., 2013; Yeomans et al., 2019). However, the literature on decision making also suggests that in certain cases where decisions require objectivity, lack of bias, and neutrality, an opposite phenomenon occurs, referred to as “algorithmic appreciation” (Sundar, 2008). This phenomenon consists of an overestimation of machine performance compared to human performance (see Castelo et al., 2019; Liu and Wei, 2019; Logg et al., 2019). The algorithmic appreciation phenomenon suggests that certain characteristics stereotypically attributed to machine performance can, under certain circumstances, elicit positive attitudes in observers. It has already been found that the use of AI led to more positive evaluations of intangible products, such as songs, compared to tangible products, such as paintings (Tigre Moura et al., 2023), but whether this might be related or not to algorithmic appreciation has not been discussed in the original paper. One relevant question is whether the appreciation of tangible, AI-generated art (i.e., paintings) might also benefit from algorithmic appreciation. For example, AI-generated paintings might be evaluated more positively in genres that require precision, accuracy, and programmed sequences of elements, and less positively in others.

Some studies have recently appeared on the different abilities of GenAI to imitate different artistic styles (Asperti et al., 2025; Ha et al., 2024; Tang et al., 2025). For example, Asperti et al. (2025) found that AI generative models appear to be more adept at imitating artistic styles such as Impressionism, Cubism, Dadaism, and Futurism, which emphasize “abstraction, bold forms, and expressive brushwork” (p. 14), whereas they face greater challenges when attempting to imitate Renaissance, Baroque, Rococo, and Naïve art styles. The inability to distinguish between human-generated and AI-generated art for some styles (reflecting gen-AI’s success at imitation) is not the same as appreciation. As we have seen in the previous pages, participants often depreciate the same artworks when they know they are AI-generated, even though they cannot distinguish them from human-made artworks. It would be interesting to study whether aversion to AI-generated art is weaker for certain styles. This aversion may be weaker for styles, such as Op Art, Cubism, and Abstract Expressionism versus Impressionism, Renaissance and Baroque Art. This hypothesis is based on the implicit assumptions participants have about the “typical style” of visual art attributed to AI, as discussed in a previous section (section 4.1). A comparative analysis of the appreciation of AI-generated art in different styles would allow us to discover various implicit assumptions.

4.2.2 Is the aversion mitigated by witnessing the production process?

Based on the few studies that have examined the impact of personal beliefs on the appreciation of AI-generated art, we know that the bias against AI-generated art is primarily driven by participants who do not attribute creative skills to AI (e.g., Agudo et al., 2022; Chamberlain et al., 2018, study 2; Di Dio et al., 2023; Hong et al., 2020), have a negative attitude towards AI (Neef et al.,

2025), and hold anthropocentric beliefs about creativity, that is, they believe creativity is a uniquely human trait—which leads them to see less creative value in, feel less awe for, and be less likely to purchase AI-generated art (Millet et al., 2023). The negative bias towards AI-generated artworks was also found to depend on personality traits (e.g., Grassini and Koivisto, 2024) and cultural differences (e.g., Wu Y. et al., 2020). All these findings highlight the importance of more frequently assessing individual factors in future studies that aim to test the appreciation of AI-generated art.

With this premise in mind, an interesting hypothesis that merits further consideration is that exposure to the production process may improve evaluations of AI-generated art. Witnessing the art production process may enhance participants’ appreciation of the “ability” required by AI to create a painting, sculpture, or a piece of music, paralleling the appreciation we have for human artists’ technical skills. This hypothesis is provisionally supported by the results of few, but interesting, studies (e.g., Chamberlain et al., 2018, study 2; Tresset and Leymarie, 2013). In particular, Chamberlain et al. (2018, study 2) demonstrated that the aesthetic response and artistic value of portraits created by robotic artists increased when participants could observe them at work. In one condition, participants were present while the robots drew portraits of individuals sitting in a chair (for full details on the robot used, see Tresset and Leymarie, 2013) and could interact with them by having their portrait drawn. In another condition, they were presented with the drawings and informed that they had been created by a robot. In a third condition, they received no information about how the drawings were created. The aesthetic ratings of the drawings were higher when participants assisted the robots while drawing as compared to both the condition where only information about the artist was provided without direct observation or interaction, and the condition where no information was given (leading participants to assume the portraits were made by a human artist).

A connected factor to consider is the anthropomorphisms that is brought into play by the process. For instance, activation of the motor and premotor cortices was found when participants were exposed to Lucio Fontana’s Cuts (Umiltà et al., 2012) or Franz Kline’s artworks—characterized by wide, marked traces of brushstroke (Sbriscia-Fioretti et al., 2013). These findings have been explained in terms of embodied simulation of the artist’s gestures during the perception of the artworks (see also Freedberg and Gallese, 2007; Oberman et al., 2007). In the context of AI, these findings raise the question of whether witnessing the art production process by AI might lead to increased engagement and a more positive evaluation of the generated output, and whether this is moderated by the amount of anthropomorphism prompted in the generative process (Waytz et al., 2014). The robots in Chamberlain et al. (2018) lacked humanoid visual characteristics, but the dynamics of their actions suggested them—particularly the robot’s alternating “looking behaviour” toward the person whose portrait was to be made and the drawing in progress. Direct exposure to the production action may activate the same mechanisms of motor simulation that occur when observing human actions.

4.2.3 Is the aversion mitigated by information about the effort and time needed to create the artwork?

We know from psychology studies, that the aesthetic appreciation of human artworks also depends on the information provided to

observers about the time the artist spent in direct contact with the artwork. The longer the contact, the higher the perceived value of the object (Newman and Bloom, 2012). Similarly, the more time and effort attributed to the artist in creating the artwork, the higher observers' ratings of liking, quality, and value of the artwork (Jucker et al., 2014, study 3; Kruger et al., 2004).³ The effect of time has been explained through effort heuristics. Assessing quality can often be challenging—for instance, determining the monetary value of a painting or the scientific contribution of a paper or book. When this is the case, people use effort as a heuristic for quality. Effort is generally a reliable indicator of quality; all else being equal, paintings that have received prolonged attention from the artist, as well as papers/books that have required extensive time to create and revise, usually result in better work. This logic supports the existence of the heuristic. However, like all heuristics, it can sometimes lead to errors. Moments of inspiration, for example, can occasionally result in unexpectedly quick and optimal outcomes.

How can we relate this to AI-generated art? If participants assume by default that AI is “quick” in doing what it does—and robots and AI systems are typically perceived to reduce effort and labor for humans (see Bechwati and Xia, 2003; Kruger et al., 2004)—this may contribute to the negative aesthetic bias toward AI-generated art. Some empirical findings support this hypothesis. For example, Horton et al. (2023) found that participants evaluated AI-labeled artworks as having taken less time to produce and as being less creative and worth less money. Magni et al. (2024, Studies 2–3) found that participants rated AI as exerting less effort than humans when performing creative tasks (e.g., designing marketing campaign posters and generating business ideas). These attributions directly correlated with negative evaluations of creativity.

Would judgments change if observers were informed that, conversely, the creation of the artwork required significant time to AI or involved the interlocking action of multiple AI networks? These are pertinent questions for empirical research. To the best of our knowledge, there is currently very little literature on this topic. Magni et al. (2024, study 4) found that in a condition where both information about the creator (AI versus human) and effort (low versus high) were manipulated, participants associated the highest creativity ratings with the human-high effort condition, followed by the human-low effort condition, the AI-high effort condition, and, finally, the AI-low effort condition. These findings indicate that effort could not substitute for the effect of the creator's identity. However, they also demonstrate an extension of effort heuristics to the domain of AI

creativity. Indeed, time and effort were not irrelevant variables for judging AI products. Bellaïche et al. (2023b) found that the artworks that participants judged to have required high effort, received higher ratings of liking and beauty when they were attributed to humans. Conversely, the artworks that participants thought required low to moderate effort received higher liking and beauty ratings when attributed to AI.

4.2.4 Is the aversion mitigated by emotional engagement with AI-generated art?

In discussing whether computers can replace human artists, Hertzmann (2018) argues that “art requires human intention, inspiration, and a desire to express something.” Various authors have emphasized that the aesthetic value of an artwork also lies in its capacity to evoke experiences charged with emotions that the creator felt and transferred into the artwork (e.g., Di Dio and Gallese, 2021; Pelowski et al., 2020, 2023). Framing the question in these terms seems to rule out any possibility of AI producing art since it is impossible for AI to transfer an experience or emotion “felt” by AI into the produced object (we can call this the *impossibility of being emotionally engaged by AI art* argument). In other words, when an artwork is created by a non-human entity, such as a computer, no human emotions can be expressed (e.g., Lu, 2005). Is this the reason why paintings labeled as AI-generated are rated by participants as less emotional than those attributed to humans (Demmer et al., 2023; Horton et al., 2023, study 1), and less awe-inspiring (Millet et al., 2023)? Or why participants exposed to identical pieces of video art or music report greater emotional arousal (“To what degree would you say that it aroused your emotion?”) when they believe the artist is human than when they believe it is AI (Agudo et al., 2022)? Or why articles written by an algorithm are considered more objective but also less emotionally engaging than those written by humans (Liu and Wei, 2019)?

On the other side, there is evidence that people can be emotionally engaged by AI art (*possibility of being emotionally engaged by AI art* argument). Bellaïche et al. (2023b) found that, for both images labeled as to human-created artworks and images labeled as to AI-created artworks, ratings of liking and beauty increased as the emotion rating (“To what extent does this artwork elicit an emotional response in you?”) increased. Furthermore, as Arielli and Manovich (2022, p. 16) point out: “The success of virtual pop stars in East Asian cultures (such as Hatsune Miku and several K-pop ‘avatar’ bands, some of them AI-driven) reveals how the public can emotionally engage with a fictional performer, follow them on social media, attend their concerts, and purchase merchandise depicting them. We could go as far as to say: fans do not love them despite, but actually because they are openly fake.”

Finally, we cannot help but notice that it is also a common experience that we sometimes appreciate a decorative pattern or the design of an object—such as a piece of furniture, a car, or a pair of shoes—without any emotional engagement with the feelings, experience, and intentions of the creator (*the indifference to emotional engagement* argument). These works can be pleasant, engaging, and entertaining in their own right.

We need further empirical evidence to clarify both the conditions that support an emotional response to AI-generated art and the necessity of an emotional engagement.

³ For example, in studies conducted by Kruger et al. (2004), participants were exposed to the same poem (in experiment 1), paintings (in experiment 2), or arms and armor (in experiment 3), but were given different information about the time the artist spent writing the poem (four or 18 h), creating the painting (four or 26 h), or making the arms and armor (15 or 110 h). The findings revealed that participants gave more favorable evaluations of the artworks when they believed the creation process took longer. The influence of time on judgment was particularly evident when it was difficult for participants to evaluate the quality of the artwork itself (achieved in their study by using low versus high-resolution images).

4.2.5 Is the aversion better understood by analysing “liking” separately in terms of pleasure and interest?

Over the past decade, several models of aesthetic liking developed in the psychology of art and empirical aesthetics have conceptualized the idea of “liking” at two different levels (e.g., Graf and Landwehr, 2015, 2017; Leder et al., 2004, 2012; Leder and Nadal, 2014; Pelowski et al., 2017), by applying dual-process theories developed in social psychology and the psychology of reasoning (e.g., Chaiken and Trope, 1999; Gawronski and Creighton, 2013; Evans, 2006, 2008; Evans and Stanovich, 2013) to the analyses of aesthetic appreciation. Here, we refer to the Pleasure and Interest model of Aesthetic liking (PIA), developed by Graf and Landwehr (2015, 2017) for human-made art, as a framework for thinking about aesthetic appreciation in AI-generated art.

According to this model, aesthetic appreciation emerges from two hierarchical, fluency-based processes. The first level is based on an automatic, default process (“gut- response”) that results in an immediate affective response of *pleasure* or *displeasure* (Reber et al., 2004, p. 365; Strack and Deutsch, 2004; Winkielman and Cacioppo, 2001; Zajonc, 1980). The second level is activated when the observer engages in further controlled and effortful processing of the object, which involves an active and reflective interaction with the stimulus (Augustin and Leder, 2006; Belke et al., 2015; Pelowski et al., 2016; Reber, 2022). The decision to process a stimulus at this second level is determined by the interplay between the observer’s motivation and the pleasure or displeasure experienced at the first level. If observers experience displeasure at the gut level (“I do not like it: why is this art?”)—or disfluency, using the model’s terminology—they may be motivated to activate the second stage in order to gain a deeper understanding of the object. This second level of processing can lead to an experience of liking that is different from the first type (pleasure) and is referred to as “interest.” Interest arises when the viewer feels that the information they have discovered or learned has improved their fluency in processing the object. Not only do they feel that they have learned something about what they are observing (the artwork), but they also experience a change in their way of thinking. They perceive themselves as adopting a different, more analytical style of processing, paying attention to non-salient attributes, and feel an increased sense of mastery as a result of this transformation. Metacognitive aspects are embedded into this second level of liking (Alter et al., 2007; Bullot and Reber, 2013; Christensen et al., 2023).

Keeping this framework in mind when analyzing human response to AI-generated art may help to identify new questions and consider new methods for inquiring whether participants like or dislike AI art. To the best of our knowledge, the PIA model has only been considered by Bellaïche et al. (2023a). Namely, it inspired them to investigate participants’ responses to art labeled as AI-generated (compared to art labeled as human-generated) at a multidimensional level, asking participants not only for ratings about liking and beauty (which capture the first level of liking, pleasure), but also for rating of profundity (“How profound is the artwork?”), meaning (“To what extent do you find this artwork personally meaningful?”), story (“To what extent can you imagine a story being communicated through this artwork?”), in addition to questions about the amount of time and effort they believed was involved in creating the artwork and the emotional response it elicited.

We believe that applying the PIA model directly to AI-generated art could inspire at least two new research directions. One direction would examine the surface-level features of AI-generated artworks that are appreciated. Do these features correspond to those appreciated in human-generated art, such as symmetry (Reber, 2002; Wurtz et al., 2008), contrast or clarity (Bornstein and D’Agostino, 1994; Halberstadt, 2006; Messinger, 1998; Reber et al., 1998; Reber et al., 2004; Song et al., 2021; van Geert and Wagemans, 2020; Winkielman et al., 2006)? Another research direction is to investigate whether exposure to information about the artist (i.e., AI), the concept behind the artwork (often encapsulated by its title), and the process leading to its creation would influence participants’ interest in the artwork, if not their perceived pleasure. Connected to this is exploring whether increased interest would also correspond to experiencing an improved sense of mastery and fluency in dealing with AI-generated art, that is, the metacognitive aspects involved in appreciation at the second level of the PIA model.

Table 2 summarizes the main biases toward AI creativity and the key research questions discussed in the previous sections.

5 Discussion

The purpose of this paper was to contribute to the ongoing discussion of how we can operationalize the constructs of creativity (which was developed to define one of the most distinctive capacities of the human mind) and aesthetic appreciation when applied to AI creations. As clarified in the Introduction, we adopted an operational perspective bridging concepts and methods taken from the literature about creativity and art appreciation as developed in cognitive psychology and empirical aesthetics into the ongoing debate on creativity and AI, with the aim to stimulate new perspectives and questions, and suggest new directions for future empirical research.

We started with two clarifications. First, by reminding that making comparisons between human-generated and AI-generated products (concerning both creative responses in general, or artworks) does not imply that the processes leading to them are the same. Second, by pointing out that previously learned knowledge is fundamental to both human and AI creativity—therefore, we should not jump too quickly to the conclusion that AI responses cannot be novel simply because they are based on prior knowledge.

We then moved on to the first of the two main topics of the paper. We examined the indices used in psychology to operationalize creativity in closed-ended tasks, where convergent creative thinking is involved, and in open-ended tasks, where divergent creative thinking is required (section entitled “Adopt indices used to assess human creativity in psychology and cognitive science to develop metrics/models for assessing AI-generated creativity”). The goals of this section were (a) to show the multidimensionality involved in the definition of creativity, (b) to provide a systematic list of indices and a clarification of the corresponding aspects they are intended to capture, and (c) to stimulate reflection on the benefits of developing a standard set of indices (and of using similar operational definitions) for research testing the performance of automated AI scoring models to evaluate the creativity of both human-generated and AI-generated responses.

The second major theme of the paper (developed in the section entitled “From empirical Aesthetics to AI”) focused on the situation

TABLE 2 Key research questions related to some “AI creativity biases”.

Biases	Research questions
Typical style: participants implicitly assume that the typical style of AI-generated images is abstract art, i.e., simple and regular compositions, bright or artificial colors, and rigid, straight, and regular shapes and lines.	<ol style="list-style-type: none">1. Is this bias rooted in prevalent exposure to early AI-generated images, and could it be corrected with extensive exposure to images resulting from the latest generation of AI algorithms?2. Does this bias reflect a broader underlying technological (precise, programmed, rigid) versus anthropological (imprecise, fantastical, flexible) divide?
Diminished value (Aversion to AI-generated art): participants tend to rate artistic performances (e.g., painting, music, and choreography) as less beautiful or aesthetically pleasing when attributed to AI rather than human creators.	<ol style="list-style-type: none">1. Is this aversion mitigated by the type of art considered? For instance, AI-generated art may be viewed more positively in genres requiring precision, accuracy, and programmed sequences of elements.2. Is this aversion mitigated by witnessing the production process? Seeing how AI creates a painting, sculpture, or piece of music may enhance participants’ appreciation of the technical skills required (likewise, as with human artists). This effect may be further moderated by the amount of anthropomorphism prompted in the generative process.3. Is this aversion mitigated by information about the effort and time needed to create the artwork (the more time and effort, the more appreciation—as for human-generated art)?4. Is the aversion mitigated by emotional engagement with AI-generated art or is emotional engagement unnecessary?5. Can this aversion be better understood by analyzing “liking” in terms of pleasure and interest separately (as foreseen by dual models of aesthetic liking)?<ul style="list-style-type: none">• Does AI-generated art spark more interest than pleasure?• Are the same stimulus-based features that are liked in human-generated art, such as symmetry, contrast, and clarity, also appreciated in AI-generated art?• Does exposure to information about the artist, the artwork, or the process of its creation enhance interest?• Does experiencing increased interest also imply that you feel more mastery and fluency in dealing with AI-generated art?

in which the creative products generated by AI are works of art, and on their aesthetic evaluation by human observers. Bridging the literature developed in psychology of art and empirical aesthetics with the literature and interest in AI, a number of new questions emerged. A first aspect focused on concerns the bias of associating abstract artworks with AI-generated art (and figurative artworks mainly with human-generated art, or both) and possible ways to empirically verify the factors underlying this bias. A second aspect concerns the diminished value usually attributed by participants to AI-generated art, compared to human-generated art. In reviewing the literature on the topic, we kept in mind some variables that have been shown in psychology to be effective in explaining participants’ aesthetic appreciation of human-generated art, and suggested a set of five questions that can be applied to the study of human appreciation of AI-generated art. These questions suggest possible future empirical research directions: Is AI-generated art aversion mitigated by the type of art considered? Is the aversion mitigated by witnessing the production process? Is it mitigated by information about the effort and time required to create the artwork? Is the aversion mitigated by emotional engagement with AI-generated art? Is it better understood by analysing “liking” separately in terms of pleasure and interest? To all these questions we hope future experimental research will find valid and intriguing answers.

Author contributions

IB: Conceptualization, Writing – original draft, Writing – review & editing, Supervision. EB: Conceptualization, Writing – original draft, Writing – review & editing. TU: Conceptualization, Writing – original draft, Writing – review & editing. RB: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors declare that the necessary funding to cover the publication fee was provided by the University of Macerata as part of the “Collaborative Constellation of Research” programme (funds attributed to Ivana Bianchi).

Acknowledgments

The idea for this work emerged from the context of the research project. "Thinking in Opposites to Improve Creativity (TOC): A Challenging Hypothesis for Cognitive Sciences and AI" (PI: Ivana Bianchi). The project is coordinated and supported by the University of Macerata under the "Collaborative Constellation of Research" program. Work partially supported by the Italian Ministry of Education and Research (MUR) in the framework of the University of Pisa FoReLab project (Departments of Excellence).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

References

- Acar, S., Berthiaume, K., Grajzel, K., Dumas, D., Flemister, C. T., and Organisciak, P. (2021). Applying automated originality scoring to the verbal form of Torrance tests of creative thinking. *Gift. Child Q.* 67, 3–17. doi: 10.1177/00169862211061
- Acar, S., Organisciak, P., and Dumas, D. (2025). Automated scoring of figural tests of creativity with computer vision. *J. Creat. Behav.* 59:e677. doi: 10.1002/jocb.677
- Agrawal, A., Gans, J., and Goldfarb, A. (2017). What to expect from artificial intelligence. Available online at: <https://mitsmr.com/2jZdf1Y> (Accessed June 13, 2025).
- Agudo, U., Arrese, M., Liberal, K. G., and Matute, H. (2022). Assessing emotion and sensitivity of AI artwork. *Front. Psychol.* 13:879088. doi: 10.3389/fpsyg.2022.879088
- Alter, A. L., Oppenheimer, D. M., Epley, N., and Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytical reasoning. *J. Exp. Psychol. Gen.* 136, 569–576. doi: 10.1037/0096-3445.136.4.569
- Amabile, T. M. (1982). Social psychology of creativity: a consensual assessment technique. *J. Pers. Soc. Psychol.* 43, 997–1013. doi: 10.1037/0022-3514.43.5.997
- Amabile, T. M. (1983). The social psychology of creativity. New York, NY: Springer.
- Amabile, T. M. (1996). Creativity in context: Update to the social psychology of creativity. Boulder, CO: Westview Press.
- Amabile, T. M., and Tighe, E. (1993). "Questions of creativity" in Creativity. ed. J. Brockman (New York, NY: Simon & Schuster), 7–27.
- Arieli, E., and Manovich, L. (2022). AI-aesthetics and the anthropocentric myth of creativity. Nodes, 1, 91–97. Available online at: <https://philpapers.org/archive/ARIAAT-6.pdf> (Accessed June 13, 2025).
- Aru, J. (2025). Artificial intelligence and the internal processes of creativity. *J. Creat. Behav.* 59:e1530. doi: 10.1002/jocb.1530
- Asperti, A., George, F., Marras, T., Stricescu, R. C., and Zanotti, F. (2025). A critical assessment of modern generative models' ability to replicate artistic styles. Available online at <https://arxiv.org/abs/2502.15856> (Accessed June 13, 2025).
- Augustin, M. D., and Leder, H. (2006). Art expertise: a study of concepts and conceptual spaces. *Psychol. Sci.* 48, 133–156. Available online at: <https://www.proquest.com/docview/212188212/fulltextPDF/B04597A8A25B4349PQ/1?accountid=15270&sourcetype=Scholarly%20Journals> (Accessed June 13, 2025).
- Baas, M., Nijstad, B. A., Boot, N. C., and De Dreu, C. K. W. (2016). Mad genius revisited: vulnerability to psychopathology, biobehavioral approach-avoidance, and creativity. *Psychol. Bull.* 142, 668–692. doi: 10.1037/bul0000049
- Baer, J., and Kaufman, J. C. (2019). "Assessing creativity with the consensual assessment technique" in The Palgrave handbook of social creativity research. eds. I. Lebeda and V. P. Glăveanu (London, UK: Palgrave Macmillan), 27–37.
- Ball, J. J., and Christensen, B. T. (2009). Analogical reasoning and mental simulation in design: two strategies linked to uncertainty resolution. *Des. Stud.* 30, 169–186. doi: 10.1016/j.destud.2008.12.005
- Ball, L. J., and Christensen, B. T. (2020). "How sticky notes support cognitive and socio-cognitive processes in the generation and exploration of creative ideas. Explorations in creativity research" in Sticky creativity: Post-it® note cognition, computers, and design. eds. B. T. Christensen, K. Halskov and C. N. Klokmoose (Cambridge, MA: Academic Press), 19–51.
- Ball, L. J., Christensen, B. T., and Halskov, K. (2021). Sticky notes as a kind of design material: how sticky notes support design cognition and design collaboration. *Des. Stud.* 76:101034. doi: 10.1016/j.destud.2021.101034
- Ball, L. J., Threadgold, E., Solowiej, A., and Marsh, J. E. (2018). Can intrinsic and extrinsic metacognitive cues shield against distraction in problem solving. *J. Cogn.* 1, 1–12. doi: 10.5334/joc.9
- Basadur, M., Graen, G. B., and Green, S. G. (1982). Training in creative problem solving: effects on ideation and problem finding and solving in an industrial research organization. *Organ. Behav. Hum. Perform.* 30, 41–70. doi: 10.1016/0030-5073(82)90233-1
- Beaty, R. E., and Johnson, D. R. (2021). Automating creativity assessment with SemDis: an open platform for computing semantic distance. *Behav. Res.* 53, 757–780. doi: 10.3758/s13428-020-01453-w
- Bechwati, N. N., and Xia, L. (2003). Do computers sweat? The impact of perceived effort of online decision aids on consumers' satisfaction with the decision process. *J. Consum. Psychol.* 13, 139–148. doi: 10.1207/S15327663JCP13-1&2_12
- Beghetto, R. A., and Kaufman, J. C. (2022). "Theories of creativity" in Creativity and innovation. Theory, research, and practice. ed. J. A. Plucker. Second ed (London, UK: Taylor and Francis), 23–36.
- Belke, B., Leder, H., and Augustin, D. (2006). Mastering style. Effects of explicit style-related information, art knowledge and affective state on appreciation of abstract paintings. *Psychol. Sci.* 48, 115–134.
- Belke, B., Leder, H., and Carbon, C.-C. (2015). When challenging art gets liked: evidences for a dual preference formation process for fluent and non-fluent portraits. *PLoS One* 10:e0131796. doi: 10.1371/journal.pone.0131796
- Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., et al. (2023b). Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. *Cogn. Res.* 8:42. doi: 10.1186/s41235-023-00499-6
- Bellaiche, L., Smith, A. P., Barr, N., Christensen, A., Williams, C., Ragnhildstveit, A., et al. (2023a). Back to the basics: abstract painting as an index of creativity. *Creat. Res. J.* 35, 698–713. doi: 10.1080/10400419.2023.2243100
- Bianchi, I., and Branchini, E. (2023). Does thinking in opposites in order to think differently improve creativity? *J. Intelligence* 11:85. doi: 10.3390/jintelligence11050085
- Bianchi, I., Branchini, E., Burro, R., Capitani, E., and Savardi, U. (2020). Overtly prompting people to "think in opposites" supports insight problem solving. *Think. Reason.* 26, 31–67. doi: 10.1080/13546783.2018.1553738
- Bianchi, I., Burro, R., Verstegen, I., Branchini, E., and Bertamini, M. (2025). Cognitive and historical information can spark interest in modern and contemporary art. *Psychol. Aesthet. Creat. Arts.* doi: 10.1037/aca0000764
- Bilić, M., Graf, M., Vaci, N., and Danek, A. H. (2021). The temporal dynamics of insight problem solving – restructuring might not always be sudden. *Think. Reason.* 27, 1–37. doi: 10.1080/13546783.2019.1705912
- Bimler, D. L., Snellock, M., and Paramei, G. V. (2019). Art expertise in construing meaning of representational and abstract artworks. *Acta Psychol.* 192, 11–22. doi: 10.1016/j.actpsy.2018.10.012
- Birdi, K. (2016). "Creativity training" in Human resource management, innovation and performance. eds. H. Shipton, P. Budhwar, P. Sparrow and A. Brown (London, UK: Palgrave Macmillan), 298–312.
- Birdi, K. (2020). Insights on impact from the development, delivery, and evaluation of the CLEAR IDEAS innovation training model. *Eur. J. Work Organ. Psychol.* 30, 400–414. doi: 10.1080/1359432X.2020.1770854
- Blair, C. S., and Mumford, M. D. (2007). Errors in idea evaluation: preference for the unoriginal? *J. Creat. Behav.* 41, 197–222. doi: 10.1002/j.2162-6057.2007.tb01288.x
- Bornstein, R. F., and D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: preliminary tests of a perceptual/attributional model of the mere exposure effect. *Soc. Cogn.* 12, 103–128. doi: 10.1521/soco.1994.12.2.103
- Branchini, E., Bianchi, I., Burro, R., Capitani, E., and Savardi, U. (2016). Can contraries prompt intuition in insight problem solving? *Front. Psychol.* 7:1962. doi: 10.3389/fpsyg.2016.01962
- Branchini, E., Capitani, E., Burro, R., Savardi, U., and Bianchi, I. (2021). Opposites in reasoning processes: do we use them more than we think, but less than we could? *Front. Psychol.* 12:715696. doi: 10.3389/fpsyg.2021.715696
- Brosowsky, N. P., Gross, M. E., Schooler, J. M., and Seli, P. (2024). "Jumping about: the role of mind-wandering and attentional flexibility in facilitating creative problem solving" in The emergence of insight. eds. C. Salvi, J. Wiley and S. M. Smith (Cambridge: Cambridge University Press), 163–180.
- Bruner, J. S. (1962). "The conditions of creativity" in Contemporary approaches to creative thinking. eds. H. Gruber, G. Terrell and M. Wertheimer (New York: Atherton), 1–30. doi: 10.1037/13117-001
- Buczak, P., Huang, H., Forthmann, B., and Doebler, P. (2023). The machines take over: a comparison of various supervised learning approaches for automated scoring of divergent thinking tasks. *J. Creat. Behav.* 57, 17–36. doi: 10.1002/jocb.559
- Bullot, N. J., and Reber, R. (2013). The artful mind meets art history: toward a psycho-historical framework for the science of art appreciation. *Behav. Brain Sci.* 36, 123–137. doi: 10.1017/S0140525X12000489
- Castelo, N., Bos, M. W., and Lehmann, D. R. (2019). Task-dependent algorithm aversion. *J. Mark. Res.* 56, 809–825. doi: 10.1177/0022243719851788
- Cetinic, E., and She, J. (2022). Understanding and creating art with AI: review and outlook. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* 18, 1–22. doi: 10.1145/3475799
- Chaiken, S., and Trope, Y. (1999). Dual-process theories in social psychology. New York, NY: Guilford Press.

- Chakrabarty, T., Laban, P., and Wu, C. S. (2025) Can AI writing be salvaged? Mitigating Idiosyncrasies and Improving Human-AI Alignment in the Writing Process through Edits. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (1–33).
- Chamberlain, R., Mullin, C., Scheerlinck, B., and Wagemans, J. (2018). Putting the art in artificial: aesthetic responses to computer-generated art. *Psychol. Aesthet. Creat. Arts* 12, 177–192. doi: 10.1037/aca0000136
- Chaudhuri, S., Dooley, M., Johnson, D., Beaty, R., and Bhattacharya, J. (2024). Evaluation of poetic creativity: predictors and the role of expertise—a multilevel approach. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000649
- Chiarella, S. G., Torromino, G., Gagliardi, D. M., Rossi, D., Babiloni, F., and Cartocci, G. (2022). Investigating the negative bias towards artificial intelligence: effects of prior assignment of AI-authorship on the aesthetic appreciation of abstract paintings. *Comput. Human Behav.* 137:107406. doi: 10.1016/j.chb.2022.107406
- Christensen, A. P., Cardillo, E. R., and Chatterjee, A. (2023). What kind of impacts can artwork have on viewers? Establishing a taxonomy for aesthetic impacts. *Brit. J. Psychol.* 114, 335–351. doi: 10.1111/bjop.12623
- Christensen, B. T., and Friis-Olivarius, M. (2020). “How do initial ideas evolve into final ones? Exploring the cognitive size, structure and life of ideas using sticky notes” in Sticky creativity: Post-it® note cognition, computers, and design. eds. B. T. Christensen, K. Halskov and C. Klokmoose (Cambridge, MA: Academic Press), 53–75.
- Cropley, D. H., and Marrone, R. L. (2025). Automated scoring of figural creativity using a convolutional neural network. *Psychol. Aesthet. Creat. Arts* 19, 77–86. doi: 10.1037/aca0000510
- Cseh, G. M., and Jeffries, K. K. (2019). A scattered CAT: a critical evaluation of the consensual assessment technique for creativity research. *Psychol. Aesthet. Creat. Arts* 13, 159–166. doi: 10.1037/aca0000220
- Dailey, L., and Mumford, M. D. (2006). Evaluative aspects of creative thought: errors in appraising the implications of new ideas. *Creat. Res. J.* 18, 385–390. doi: 10.1207/s15326934crj1803_11
- Danek, A. H. (2018). “Magic tricks, sudden restructuring and the Aha! Experience: a new model of non-monotonic problem solving” in Insight: On the origins of new ideas. ed. F. Vallée-Tourangeau (London, UK: Routledge), 51–78.
- Danek, A. H., and Flanagan, V. L. (2019). Cognitive conflict and restructuring: the neural basis of two core components of insight. *AIMS Neurosci.* 6, 60–84. doi: 10.3934/Neuroscience.2019.2.60
- Danek, A. H., Öllinger, M., Fraps, T., Grothe, B., and Flanagan, V. L. (2015). An fMRI investigation of expectation violation in magic tricks. *Front. Psychol.* 6:84. doi: 10.3389/fpsyg.2015.00084
- Danek, A. H., and Salvi, C. (2020). Moment of truth: why Aha! Experiences are correct. *J. Creat. Behav.* 54, 484–486. doi: 10.1002/jocb.380
- Danek, A. H., and Wiley, J. (2017). What about false insights? Deconstructing the Aha! Experience along its multiple dimensions for correct and incorrect solutions separately. *Front. Psychol.* 7:2077. doi: 10.3389/fpsyg.2016.02077
- Danek, A. H., Williams, J., and Wiley, J. (2020). Closing the gap: connecting sudden representational change to the subjective Aha! Experience in insightful problem solving. *Psychol. Res.* 84, 111–119. doi: 10.1007/s00426-018-0977-8
- Darda, K. M., and Cross, E. S. (2023). The computer, a choreographer? Aesthetic responses to randomly-generated dance choreography by a computer. *Heliyon*. 9:e12750. doi: 10.1016/j.heliyon.2022.e12750
- Demmer, T. R., Kühnapfel, C., Fingerhut, J., and Pelowski, M. (2023). Does an emotional connection to art really require a human artist? Emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. *Comput Human Behav.* 148:107875. doi: 10.1016/j.chb.2023.107875
- DeYoung, C. G., Flanders, J. L., and Peterson, J. B. (2008). Cognitive abilities involved in insight problem solving: an individual differences model. *Creat. Res. J.* 20, 278–290. doi: 10.1080/10400410802278719
- Di Dio, C., Ardizzi, M., Schieppati, S. V., Massaro, D., Gilli, G., Gallese, V., et al. (2023). Art made by artificial intelligence: the effect of authorship on aesthetic judgments. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000602
- Di Dio, C., and Gallese, V. (2021). “Moving toward emotions in the aesthetic experience” in Brain, beauty, and art: essays bringing neuroaesthetics into focus. eds. A. Chatterjee and E. Cardillo (Oxford: Oxford University Press), 682–687.
- Diedrich, J., Benedek, M., Jauk, E., and Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychol. Aesthet. Creat. Arts* 9, 35–40. doi: 10.1037/a0038688
- Dietvorst, B. J., Simmons, J. P., and Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* 144, 114–126. doi: 10.1037/xge0000033
- Du Sautoy, M. (2020). The creativity code: Art and innovation in the age of AI. Harvard: Harvard University Press.
- Dumas, D., Organisciak, P., and Doherty, M. (2021). Measuring divergent thinking originality with human raters and text-mining models: a psychometric comparison of methods. *Psychol. Aesthet. Creat. Arts* 15, 645–663. doi: 10.1037/aca0000319
- Elgammal, A. (2019). AI is blurring the definition of artist: advanced algorithms are using machine learning to create art autonomously. *Am. Sci.* 107, 18–22. Available online at: <https://www.jstor.org/stable/26586036>
- Elgammal, A., Liu, B., Elhoseiny, M., and Mazzone, M. (2017). CAN: creative adversarial networks, generating “art” by learning about styles and deviating from style norms. Available online at: <https://arxiv.org/abs/1706.07068> (Accessed June 15, 2025).
- Ellamil, M., Dobson, C., Beeman, M., and Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage* 59, 1783–1794. doi: 10.1016/j.neuroimage.2011.08.008
- Ellis, J. J., Glaholt, M. G., and Reingold, E. M. (2011). Eye movements reveal solution knowledge prior to insight. *Conscious. Cogn.* 20, 768–776. doi: 10.1016/j.concog.2010.12.007
- Ellis, J. J., and Reingold, E. M. (2014). The Einstellung effect in anagram problem solving: evidence from eye movements. *Front. Psychol.* 5:679. doi: 10.3389/fpsyg.2014.00679
- Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: extension and evaluation. *Psychon. Bull. Rev.* 13, 378–395. doi: 10.3758/BF03193858
- Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629
- Evans, J. S. B. T., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241. doi: 10.1177/1745691612460685
- Fedor, A., Szathmáry, E., and Öllinger, M. (2015). Problem solving stages in the five square problem. *Front. Psychol.* 6:1050. doi: 10.3389/fpsyg.2015.01050
- Finke, R. A., Ward, T. B., and Smith, S. M. (1992). Creative cognition: Theory, research and applications. Cambridge (MA): MIT Press.
- Fischinger, T., Kaufman, M., and Schlotz, W. (2020). If it's Mozart, it must be good? The influence of textual information and age on musical appreciation. *Psychol. Music* 48, 579–597. doi: 10.1177/0305735618812216
- Ford, C. M., and Gioia, D. A. (2000). Factors influencing creativity in the domain of managerial decision making. *J. Manage.* 26, 705–732. doi: 10.1016/S0149-2063(00)00053-2
- Forthmann, B., Holling, H., Zandi, N., Gerwig, A., Çelik, P., Storme, M., et al. (2017). Missing creativity: the effect of cognitive workload on rater (dis-)agreement in subjective divergent-thinking scores. *Think. Skills* 23, 129–139. doi: 10.1016/j.tsc.2016.12.005
- Forthmann, B., Szardenings, C., and Holling, H. (2020). Understanding the confounding effect of fluency in divergent thinking scores: revisiting average scores to quantify artifactual correlation. *Psychol. Aesthet. Creat. Arts* 14, 94–112. doi: 10.1037/aca0000196
- Freedberg, D., and Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends Cogn. Sci.* 11, 197–203. doi: 10.1016/j.tics.2007.02.003
- Fuchs, C., Schreier, M., and van Osselaer, S. M. J. (2015). The handmade effect: what's love got to do with it? *J. Mark.* 79, 98–110. doi: 10.1509/jm.14.0018
- Gangadharbatla, H. (2022). The role of AI attribution knowledge in the evaluation of artwork. *Empir. Stud. Arts* 40, 125–142. doi: 10.1177/0276237421994697
- Garcia, M. B. (2024). The paradox of artificial creativity: challenges and opportunities of generative AI artistry. *Creat. Res. J.* 8, 1–14. doi: 10.1080/10400419.2024.2354622
- Gawronski, B., and Creighton, L. A. (2013). “Dual process theories” in The Oxford handbook of social cognition. ed. D. E. Carlston (Washington, DC: American Psychologist Association), 282–312.
- Gilhooly, K. J., Georgiou, G. J., Sirota, M., and Paphiti-Galeano, A. (2019). “Incubation and suppression processes in creative problem solving” in Insight and creativity in problem solving. eds. K. J. Gilhooly, L. J. Ball and L. Macchi (London, UK: Routledge), 130–146.
- Gilhooly, K. J., and Murphy, P. (2005). Differentiating insight from non-insight problems. *Think Reason.* 11, 279–302. doi: 10.1080/1354678042000187
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2020). Generative adversarial networks. *Commun. ACM* 63, 139–144. doi: 10.1145/3422622
- Graf, M., Danek, A. H., Vaci, N., and Bilalić, M. (2023). Tracing cognitive processes in insight problem solving: using GAMs and change point analysis to uncover restructuring. *J. Intelligence* 11:86. doi: 10.3390/jintelligence11050086
- Graf, L. K. M., and Landwehr, J. R. (2015). A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking. *Personal. Soc. Psychol. Rev.* 19, 395–410. doi: 10.1177/1088868315574978
- Graf, L. K. M., and Landwehr, J. R. (2017). Aesthetic pleasure versus aesthetic interest: the two routes to aesthetic liking. *Front. Psychol.* 8:15. doi: 10.3389/fpsyg.2017.00015
- Grassini, S., and Koivisto, M. (2024). Understanding how personality traits, experiences, and attitudes shape negative bias toward AI-generated artworks. *Sci. Rep.* 14:4113. doi: 10.1038/s41598-024-54294-4
- Grassini, S., and Koivisto, N. (2025). Artificial creativity? Evaluating AI against human performance in creative interpretation of visual stimuli. *Int. J. Hum. Comput. Interact.* 41, 4037–4048. doi: 10.1080/10447318.2024.2345430
- Green, O. (2020). Bob the robot: Exploring the universe - a cozy bedtime story produced by artificial intelligence. Stockholm: Kindle Direct Publishing.

- Guilford, J. P. (1967a). Creativity: yesterday, today, and tomorrow. *J. Creat. Behav.* 1, 3–14. doi: 10.1002/j.2162-6057.1967.tb00002.x
- Guilford, J. P. (1967b). The nature of human intelligence. New York: McGraw-Hill.
- Guzik, E. E., Byrge, C., and Gilde, C. (2023). The originality of machines: AI takes the Torrance test. *J. Creat.* 33:100065. doi: 10.1016/j.jyoc.2023.100065
- Ha, A. Y. J., Passananti, J., Bhaskar, R., Shan, S., Southen, R., Zheng, H., et al. (2024). Organic or diffused: Can we distinguish human art from ai-generated images?. CCS'24: Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security, pp. 4822–4836.
- Halberstadt, J. (2006). The generality and ultimate origins of the attractiveness of prototypes. *Personal. Soc. Psychol. Rev.* 10, 166–183. doi: 10.1207/s15327957pspr1002_5
- Hao, N., Ku, Y., Liu, M., Hu, Y., Bodner, M., Grabner, R. H., et al. (2016). Reflection enhances creativity: beneficial effects of idea evaluation on idea generation. *Brain Cogn.* 103, 30–37. doi: 10.1016/j.bandc.2016.01.005
- Hass, R. W. (2017). Semantic search during divergent thinking. *Cognition* 166, 344–357. doi: 10.1016/j.cognition.2017.05.039
- Hertzmann, A. (2018). Can computers create art? *Art* 7:18. doi: 10.3390/arts7020018
- Hertzmann, A. (2020). Computers do not make art, people do. *Commun. ACM* 63, 45–48. doi: 10.1145/3347092
- Hocevar, D. (1979). A comparison of statistical infrequency and subjective judgment as criteria in the measurement of originality. *J. Pers. Assess.* 43, 297–299. doi: 10.1207/s15327752jpa4303_13
- Hodge, S. (2012). Why your five year old could not have done that: Modern art explained. London, UK: Thames & Hudson.
- Hong, J. W., and Curran, N. M. (2019). Artificial intelligence, artists, and art: attitudes toward artwork produced by humans vs. artificial intelligence. *ACM Trans. Multimed. Comput. Commun. Appl.* 15:58, 1–16. doi: 10.1145/3326337
- Hong, J. W., Fischer, K., Ha, Y., and Zeng, Y. (2022). Human, I wrote a song for you: an experiment testing the influence of machines' attributes on the AI-composed music evaluation. *Comput. Hum. Behav.* 131:107239. doi: 10.1016/j.chb.2022.107239
- Hong, J. W., Peng, Q., and Williams, D. (2020). Are you ready for artificial Mozart and Skrillex? An experiment testing expectancy violation theory and AI music. *New Media Soc.* 23, 1920–1935. doi: 10.1177/1461444820925798
- Horton, C. B., White, M. W., and Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Sci. Rep.* 13:19001. doi: 10.1038/s41598-023-45202-3
- Hubert, K. F., Awa, K. N., and Zabelina, D. L. (2024). The current state of artificial intelligence generative language models is more creative than humans on divergent thinking tasks. *Sci. Rep.* 14:3440. doi: 10.1038/s41598-024-53303-w
- Huo, K. (2020). Performance incentives, divergent thinking training, and creative problem solving. *J. Manage. Account. Res.* 32, 159–176. doi: 10.2308/jmar-52479
- Ismayilzade, M., Stevenson, C., and van der Plas, L. (2025). Evaluating creative short story generation in humans and large language models. Available online at: <https://arxiv.org/abs/2411.02316> (Accessed June 15, 2025)
- Ivancovsky, T., Shamay-Tsoory, S., Lee, J., Morio, H., and Kurman, J. (2019). A dual process model of generation and evaluation: a theoretical framework to examine cross-cultural differences in the creative process. *Pers. Individ. Differ.* 139, 60–68. doi: 10.1016/j.paid.2018.11.012
- Jones-Chick, R., Kelloway, E. K., and Birdi, K. (2022). CLEAR IDEAS: can idea implementation training enhance the development of new ideas beyond idea generation training? *J. Creat. Behav.* 56, 79–91. doi: 10.1002/jocb.517
- Jucker, J. L., Barrett, J. L., and Wlodarski, R. (2014). "I just don't get it": perceived artists' intentions affect art evaluations. *Empir. Stud. Arts* 32, 149–182. doi: 10.2190/EM.32.2.c
- Karimi, P., Grace, K., Maher, M. L., and Davis, N. (2018). Evaluating creativity in computational co-creative systems. Available online at: <https://arxiv.org/abs/1807.09886> (Accessed June 15, 2025)
- Kaufman, J. C., Baer, J., Cole, J. C., and Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creat. Res. J.* 20, 171–178. doi: 10.1080/10400410802059929
- Kaufman, J. C., Lee, J., Baer, J., and Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: new evidence of validity. *Think. Skills* 2, 96–106. doi: 10.1016/j.tsc.2007.04.002
- Kaufman, J. C., and Sternberg, R. J. (Eds.) (2010). The Cambridge handbook of creativity. Cambridge: Cambridge University Press.
- Kenett, Y. N. (2019). What can quantitative measures of semantic distance tell us about creativity? *Curr. Opin. Behav. Sci.* 27, 11–16. doi: 10.1016/j.cobeha.2018.08.010
- Kenett, Y. N., and Faust, M. (2019). A semantic network cartography of the creative mind. *Trends Cogn. Sci.* 23, 271–274. doi: 10.1016/j.tics.2019.01.007
- Kern, F. B., Wu, C. T., and Chao, Z. C. (2024). Assessing novelty, feasibility and value of creative ideas with an unsupervised approach using GPT-4. *Brit. J. Psychol.* 1–20. doi: 10.1111/bjop.12720
- Kienitz, E., Quintin, E.-M., Saggat, M., Bott, N. T., Royalty, A., Wei-Chen Hong, D., et al. (2014). Targeted intervention to increase creative capacity and performance: a randomized controlled pilot study. *Think Skills* 13, 57–66. doi: 10.1016/j.tsc.2014.03.002
- Kirk, U., Skov, M., Hulme, O., Christensen, M. S., and Zeki, S. (2009). Modulation of aesthetic value by semantic context: an fMRI study. *NeuroImage* 44, 1125–1132. doi: 10.1016/j.neuroimage.2008.10.009
- Knapp, R. H., and Wulff, A. (1963). Preferences for abstract and representational art. *J. Soc. Psychol.* 60, 255–262. doi: 10.1080/0022454.1963.9922196
- Knoblich, G., Ohlsson, S., and Raney, G. E. (2001). An eye movement study of insight problem solving. *Mem Cognit.* 29, 1000–1009. doi: 10.3758/BF03195762
- Köbis, N., and Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Comput. Hum. Behav.* 114:106553. doi: 10.1016/j.chb.2020.106553
- Koivisto, M., and Grassini, S. (2023). Best humans still outperform artificial intelligence in a creative divergent thinking task. *Sci. Rep.* 13:13601. doi: 10.1038/s41598-023-40858-3
- Kruger, J., Wirtz, D., Van Boven, L., and Altermatt, T. W. (2004). The effort heuristic. *J. Exp. Soc. Psychol.* 40, 91–98. doi: 10.1016/S0022-1031(03)00065-9
- Landers, M. (2025). Modern Prometheus: tracing the ill-defined path to AGI. *AI & Soc.* 11, 1–19. doi: 10.1007/s00146-025-02363-1
- Laukkonen, R. E., Ingledew, D. J., Grimmer, H. J., Schooler, J. W., and Tangen, J. M. (2021). Getting a grip on insight: real-time and embodied Aha experiences predict correct solutions. *Cognit. Emot.* 35, 918–935. doi: 10.1080/02699931.2021.1908230
- Leder, H., Belke, B., Oeberst, A., and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *Brit. J. Psychol.* 95, 489–508. doi: 10.1348/0007126042369811
- Leder, H., Carbon, C. C., and Ripsas, A. (2006). Entitling art: influence of title information on understanding and appreciation of paintings. *Acta Psychol.* 121, 176–198. doi: 10.1016/j.actpsy.2005.08.005
- Leder, H., Gerger, G., Dressler, S. G., and Schabmann, A. (2012). How art is appreciated. *Psychol. Aesthet. Creat. Arts* 6, 2–10. doi: 10.1037/a0026396
- Leder, H., and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode - developments and challenges in empirical aesthetics. *Brit. J. Psychol.* 105, 443–464. doi: 10.1111/bjop.12084
- Litchfield, R. C., Gilson, L. L., and Gilson, P. W. (2015). Defining creative ideas: toward a more nuanced approach. *Group Organ Manage* 40, 238–265. doi: 10.1177/1059601115574945
- Liu, P., Chu, Y., Zhao, Y., and Zhai, S. (2025). Machine creativity: aversion, appreciation, or indifference? *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000739
- Liu, B., and Wei, L. (2019). Machine authorship in situ: effect of news organization and news genre on news credibility. *Digit. Journal.* 7, 635–657. doi: 10.1080/21670811.2018.1510740
- Logg, J. M., Minson, J. A., and Moore, D. A. (2019). Algorithm appreciation: people prefer algorithmic to human judgment. *Organ. Behav. Hum. Decis. Process.* 151, 90–103. doi: 10.1016/j.obhdp.2018.12.005
- Long, H., and Pang, W. (2015). Rater effects in creativity assessment: a mixed methods investigation. *Think. Skills Creat.* 15, 13–25. doi: 10.1016/j.tsc.2014.10.004
- Long, H., and Wang, J. (2022). Dissecting reliability and validity evidence of subjective creativity assessment: a literature review. *Educ. Psychol. Rev.* 34, 1399–1443. doi: 10.1007/s10648-022-09679-0
- Lu, L. F. L. (2005). Pre-service art teacher negative attitudes and perceptions of computer-generated art imagery: recommendations for pre-service art education programs. *Vis. Arts. Res.* 31, 89–102.
- Lucas, B. J., and Nordgren, L. F. (2020). The creative cliff illusion. *Proc. Natl. Acad. Sci. USA* 117, 19830–19836. doi: 10.1073/pnas.2005620117
- Luchini, S. A., Maliakkal, N. T., DiStefano, P. V., Laverghetta, A. Jr., Patterson, J. D., Beaty, R. E., et al. (2025). Automated scoring of creative problem solving with large language models: A comparison of originality and quality ratings. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000736
- Magni, F., Park, J., and Chao, M. M. (2024). Humans as creativity gatekeepers: are we biased against AI creativity? *J. Bus. Psychol.* 39, 643–656. doi: 10.1007/s10869-023-09910-x
- Maier, N. R. F. (1930). Reasoning in humans. I. On direction. *J. Comp. Psychol.* 10, 115–143. doi: 10.1037/h0073232
- Manovich, L. (2022). AI and myths of creativity. *Archit. Des.* 92, 60–65. doi: 10.1002/ad.2814
- Manu, A. (2024). Transcending imagination: Artificial intelligence and the future of creativity. Boca Raton, FL/Abington Oxon, UK: CRC Press.
- Marsh, J. E., Threadgold, E., Barker, M. E., Litchfield, M., Degno, F., and Ball, L. J. (2021). The susceptibility of compound remote associate problems to disruption by irrelevant sound: a window onto the component processes underpinning creative cognition? *J. Cogn. Psychol.* 33, 793–822. doi: 10.1080/20445911.2021.1900201

- Mastandrea, S., Bartoli, G., and Carrus, G. (2011). The automatic aesthetic evaluation of different art and architectural styles. *Psychol. Aesthet. Creat. Arts* 5, 126–134. doi: 10.1037/a0021126
- Mastandrea, S., Wagoner, J. A., and Hogg, M. A. (2021). Liking for abstract and representational art: national identity as an art appreciation heuristic. *Psychol. Aesthet. Creat. Arts* 15, 241–249. doi: 10.1037/aca0000272
- Mastria, S., Agnoli, S., and Corazza, G. E. (2019). How does emotion influence the creativity evaluation of exogenous alternative ideas? *PLoS One* 14:e0219298. doi: 10.1371/journal.pone.0219298
- Mazzone, M., and Elgammal, A. (2019). Art, creativity, and the potential of artificial intelligence. *Art* 8:26. doi: 10.3390/arts8010026
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychol. Rev.* 69, 220–232. doi: 10.1037/h0048850
- Mednick, S. A., and Mednick, M. T. (1967). Manual for the remote associates test: Form 1. Boston, Mass: Houghton-Mifflin.
- Messer, U. (2024). Co-creating art with generative artificial intelligence: implications for artworks and artists. *Comput. Hum. Behav.* 2:100056. doi: 10.1016/j.chbah.2024.100056
- Messinger, S. M. (1998). Pleasure and complexity: Berlyne revisited. *J. Psychol.* 132, 558–560. doi: 10.1080/00223989809599288
- Messingschlager, T., and Appel, M. (2025). Mind ascribed to AI and the appreciation of AI-generated art. *New Media Soc.* 27, 1673–1692. doi: 10.1177/14614448231200248
- Meyer, B. J. E., and Pollard, C. K. (2006). “Applied learning and aging: A closer look at reading” in *Handbook of the psychology of aging*. eds. J. E. Birren and K. W. Schaie. Sixth ed (Cambridge, MA: Academic Press), 233–260.
- Mikalonytė, E. S., and Kneer, M. (2022). Can artificial intelligence make art?: folk intuitions as to whether AI-driven robots can be viewed as artists and produce art. *ACM Trans. Hum. Robot Interact.* 11:43. doi: 10.1145/3530875
- Millet, K., Buehler, F., Du, G., and Kokkoris, M. (2023). Defending humankind: anthropocentric bias in the appreciation of AI art. *Comput. Hum. Behav.* 143:107707. doi: 10.1016/j.chb.2023.107707
- Millis, K. (2001). Making meaning brings pleasure: the influence of titles on aesthetic experiences. *Emotion* 1, 320–329. doi: 10.1037/1528-3542.1.3.320
- Moffat, D. C., and Kelly, M. (2006). An investigation into people's bias against computational creativity in music composition. *Assessment* 13, 1–8.
- Moreau, C. P., Lehmann, D. R., and Markman, A. B. (2001). Entrenched knowledge structures and consumer responses to new products. *J. Mark. Res.* 38, 14–29. doi: 10.1509/jmkr.38.1.14.18836
- Moura, F. T. (2024). “Rethinking creativity frameworks for artificial intelligence” in *Artificial intelligence, co-creation and creativity*. ed. F. T. Moura (London, UK: Routledge), 32–44.
- Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creat. Res. J.* 15, 107–120. doi: 10.1080/10400419.2003.9651403
- Myszkowski, N., and Storme, M. (2019). Judge response theory? A call to upgrade our psychometrical account of creativity judgments. *Psychol. Aesthet. Creat. Arts* 13, 167–175. doi: 10.1037/aca0000225
- Neef, N. E., Zabel, S., Papoli, M., and Otto, S. (2025). Drawing the full picture on diverging findings: adjusting the view on the perception of art created by artificial intelligence. *AI & Soc.* 40, 2859–2879. doi: 10.1007/s00146-024-02020-z
- Newman, G. E., and Bloom, P. (2012). Art and authenticity: the importance of originals in judgments of value. *J. Exp. Psychol. Gen.* 141, 558–569. doi: 10.1037/a0026035
- Newton, D. (2019). Generative deep learning in architectural design. *TAD* 3, 176–189. doi: 10.1080/24751448.2019.1640536
- Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., and Baas, M. (2010). The dual pathway to creativity model: creative ideation as a function of flexibility and persistence. *Eur. Rev. Soc. Psychol.* 21, 34–77. doi: 10.1080/10463281003765323
- Noll, A. M. (1967). The digital computer as a creative medium. *IEEE Spectr.* 4, 89–95. doi: 10.1109/MSPEC.1967.5217127
- Oberman, L. M., McKeery, J. P., Ramachandran, V. S., and Pineda, J. A. (2007). EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. *Neurocomputing* 70, 2194–2203. doi: 10.1016/j.neucom.2006.02.024
- Öllinger, M., Fedor, A., Brodt, S., and Szathmáry, E. (2017). Insight into the ten-penny problem: guiding search by constraints and maximization. *Psychol. Res.* 81, 925–938. doi: 10.1007/s00426-016-0800-3
- Olson, J. A., Nahas, J., Chmoulevitch, D., Cropper, S. J., and Webb, M. E. (2021). Naming unrelated words predicts creativity. *PNAS* 118:e2022340118. doi: 10.1073/pnas.2022340118
- Orwig, W., Edenbaum, E. R., Greene, J. D., and Schacter, D. (2024). The language of creativity: evidence from humans and large language models. *J. Creat. Behav.* 58, 128–136. doi: 10.1002/jocb.636
- Patterson, J. D., Barbot, B., Lloyd-Cox, J., and Beaty, R. E. (2024). AuDrA: an automated drawing assessment platform for evaluating creativity. *Behav. Res.* 56, 3619–3636. doi: 10.3758/s13428-023-02258-3
- Pelowski, M., Markey, P. S., Forster, M., Gerger, G., and Leder, H. (2017). Move me, astonish me... Delight my eyes and brain: the Vienna integrated model of top-down and bottom-up processes in art perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Phys Life Rev* 21, 80–125. doi: 10.1016/j.plrev.2017.02.003
- Pelowski, M., Markey, P. S., Luring, J. O., and Leder, H. (2016). Visualizing the impact of art: an update and comparison of current psychological models of art experience. *Front. Hum. Neurosci.* 10:160. doi: 10.3389/fnhum.2016.00160
- Pelowski, M., Specker, E., Boddy, J., Immelmann, B., Haiduk, F., Spezie, G., et al. (2023). Together in the dark? Investigating the understanding and feeling of intended emotions between viewers and professional artists at the Venice biennale. *Psychol. Aesthet. Creat. Arts* 17, 772–792. doi: 10.1037/aca0000436
- Pelowski, M., Specker, E., Gerger, G., Leder, H., and Weingarden, L. S. (2020). Do you feel like I do? A study of spontaneous and deliberate emotion sharing and understanding between artists and perceivers of installation art. *Psychol. Aesthet. Creat. Arts* 14, 276–293. doi: 10.1037/aca0000201
- Pihko, E., Virtanen, A., Saarinen, V.-M., Pannasch, S., Hirvenkari, L., Tossavainen, T., et al. (2011). Experiencing art: the influence of expertise and painting abstraction level. *Front. Hum. Neurosci.* 5:94. doi: 10.3389/fnhum.2011.00094
- Plucker, J., Makel, M., and Qian, M. (2019). “Assessment of creativity” in *The Cambridge handbook of creativity*. eds. J. Kaufman and R. Sternberg (Cambridge: Cambridge University Press), 44–68.
- Poetz, M. K., and Schreier, M. (2012). The value of crowdsourcing: can users really compete with professionals in generating new product ideas? *J. Prod. Innovat. Manag.* 29, 245–256. doi: 10.1111/j.1540-5885.2011.00893.x
- Prabhakaran, R., Green, A. E., and Gray, J. R. (2014). Thin slices of creativity: using single-word utterances to assess creative cognition. *Behav. Res.* 46, 641–659. doi: 10.3758/s13428-013-0401-7
- Primi, R. (2014). Divergent productions of metaphors: combining many-facet Rasch measurement and cognitive psychology in the assessment of creativity. *Psychol. Aesthet. Creat. Arts* 8, 461–474. doi: 10.1037/a0038055
- Primi, R., Silvia, P. J., Jauk, E., and Benedek, M. (2019). Applying many-facet Rasch modeling in the assessment of creativity. *Psychol. Aesthet. Creat. Arts* 13, 176–186. doi: 10.1037/aca0000230
- Putman, V. L., and Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. *J. Creat. Behav.* 43, 23–39. doi: 10.1002/j.2162-6057.2009.tb01304.x
- Ragot, M., Martin, N., and Cojean, S. (2020). AI-generated vs. human artworks. A perception bias towards artificial intelligence? CHI EA '20: Extended abstracts of the 2020 CHI conference on human factors in computing systems. Available online at: <https://dl.acm.org/doi/abs/10.1145/3334480.3382892> (Accessed June 15, 2025).
- Reber, R. (2002). Reasons for the preference for symmetry. *Behav. Brain Sci.* 25, 415–416. doi: 10.1017/S0140525X02350076
- Reber, R. (2022). “Appreciation modes in empirical aesthetics” in *The Oxford handbook of empirical aesthetics*. eds. M. Nadal and O. Vartanian (Oxford: Oxford University Press), 116–135.
- Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personal. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3
- Reber, R., Winkielman, P., and Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychol. Sci.* 9, 45–48. doi: 10.1111/1467-9280.00008
- Redaelli, R. (2025). Intentionality gap and preter-intentionality in generative artificial intelligence. *AI & Soc.* 40, 2525–2532. doi: 10.1007/s00146-024-02007-w
- Reiter-Palmon, R., Forthmann, B., and Barbot, B. (2019). Scoring divergent thinking tests: a review and systematic framework. *Psychol. Aesthet. Creat. Arts* 13, 144–152. doi: 10.1037/aca0000227
- Rietzschel, E. F., Nijstad, B. A., and Stroebe, W. (2006). Productivity is not enough: a comparison of interactive and nominal brainstorming groups on idea generation and selection. *J. Exp. Soc. Psychol.* 42, 244–251. doi: 10.1016/j.jesp.2005.04.005
- Rietzschel, E. F., Nijstad, B. A., and Stroebe, W. (2010). The selection of creative ideas after individual idea generation: choosing between creativity and impact. *Brit. J. Psychol.* 101, 47–68. doi: 10.1348/000712609X414204
- Rietzschel, E. F., Nijstad, B. A., and Stroebe, W. (2019). “Why great ideas are often overlooked: a review and theoretical analysis of research on idea evaluation and selection” in *Team creativity and innovation*. ed. R. Reiter-Palmon (Oxford: Oxford University Press), 179–197.
- Ritter, S. M., Gu, X., Crijns, M., and Biekens, P. (2020). Fostering students' creative thinking skills by means of a one-year creativity training program. *PLoS One* 15:e0229773. doi: 10.1371/journal.pone.0229773
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Available online at: https://openaccess.thecvf.com/content/CVPR2022/papers/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR2022_paper.pdf (Accessed June 15, 2025).
- Runco, M. A. (Ed.) (2012). *Creativity research handbook*. New York: Hampton Press.

- Runco, M. A. (2023). AI can only produce artificial creativity. *J. Creat.* 33:100063. doi: 10.1016/j.jyoc.2023.100063
- Runco, M. A. (2025). Updating the standard definition of creativity to account for the artificial creativity of AI. *Creat. Res. J.* 37, 1–5. doi: 10.1080/10400419.2023.2257977
- Runco, M. A., and Jaeger, G. J. (2012). The standard definition of creativity. *Creat. Res. J.* 24, 92–96. doi: 10.1080/10400419.2012.650092
- Russell, P. A. (2003). Effort after meaning and the hedonic value of paintings. *Brit. J. Psychol.* 94, 99–110. doi: 10.1348/000712603762842138
- Salvi, C., and Bowden, E. (2024). “Insight: what happens backstage?” in *The emergence of insight*. eds. C. Salvi, J. Wiley and S. M. Smith (Cambridge: Cambridge University Press), 280–308.
- Salvi, C., Bricolo, E., Kounios, J., Bowden, E., and Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Think Reason.* 22, 443–460. doi: 10.1080/13546783.2016.1141798
- Samo, A., and Highhouse, S. (2023). Artificial intelligence and art: identifying the aesthetic judgment factors that distinguish human- and machine-generated artwork. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000570
- Sbriscia-Fioretti, B., Berchio, C., Freedberg, D., Gallese, V., and Umiltà, M. A. (2013). ERP modulation during observation of abstract paintings by Franz Kline. *PLoS One* 8:e75241. doi: 10.1371/journal.pone.0075241
- Schröter, J. (2019). “Artificial intelligence and the democratization of art” in *The democratization of artificial intelligence: Net politics in the era of learning algorithms*. ed. A. Sudmann (Bielefeld: Transcript Verlag), 297–312.
- Seli, P., Ragnhildstveit, A., Orwig, W., Bellaiche, L., Spooner, S., and Barr, N. (2025). Beyond the brush: human versus artificial intelligence creativity in the realm of generative art. *Psychol. Aesthet. Creat. Arts*, 1–9. doi: 10.1037/aca0000743
- Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., and Medow, M. A. (2013). Why do patients derogate physicians who use a computer-based diagnostic support system? *Med. Decis. Mak.* 33, 108–118. doi: 10.1177/0272989X12453501
- Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K., and Belfi, A. M. (2023). AI composer bias: listeners like music less when they think it was composed by an AI. *J. Exp. Psychol. Appl.* 29, 676–692. doi: 10.1037/xap0000447
- Silvia, P. J. (2008). Discernment and creativity: how well can people identify their most creative ideas? *Psychol. Aesthet. Creat. Arts* 2, 139–146. doi: 10.1037/1931-3896.2.3.139
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., et al. (2008). Assessing creativity with divergent thinking tasks: exploring the reliability and validity of new subjective scoring methods. *Psychol. Aesthet. Creat. Arts* 2, 68–85. doi: 10.1037/1931-3896.2.2.68
- Simon, H. A. (2001). Creativity in the arts and the sciences. *Kenyon Rev.* 23, 203–220. Available online at: <https://www.jstor.org/stable/4338222>
- Simonton, D. K. (2004). Creativity in science: Chance, logic, genius, and zeitgeist. Cambridge: Cambridge University Press.
- Simonton, D. K. (2014). More method in the mad-genius controversy: a historiometric study of 204 historic creators. *Psychol. Aesthet. Creat. Arts* 8, 53–61. doi: 10.1037/a0035367
- Sio, U. N., Kotovsky, K., and Cagan, J. (2018). Silence is golden: the effect of verbalization on group performance. *J. Exp. Psychol. Gen.* 147, 939–944. doi: 10.1037/xge0000456
- Sio, U. N., and Ormerod, T. C. (2015). Incubation and cueing effects in problem-solving: set aside the difficult problems but focus on the easy ones. *Think Reason.* 21, 113–129. doi: 10.1080/13546783.2014.886626
- Snyder, H. T., Hammond, J. A., Grohman, M. G., and Katz-Buonincontro, J. (2019). Creativity measurement in undergraduate students from 1984–2013: a systematic review. *Psychol. Aesthet. Creat. Arts* 13, 133–143. doi: 10.1037/aca0000228
- Song, J., Kwak, Y., and Kim, C. Y. (2021). Familiarity and novelty in aesthetic preference: the effects of the properties of the artwork and the beholder. *Front. Psychol.* 12:694927. doi: 10.3389/fpsyg.2021.694927
- Sovhyra, T. (2021). Artificial intelligence and issue of authorship and uniqueness for works of art (technological research of the next Rembrandt). *Culture Arts Modern World* 22, 156–163. doi: 10.31866/2410-1915.22.2021.235903
- Stein, M. I. (1953). Creativity and culture. *J. Psychol.* 36, 311–322. doi: 10.1080/00223980.1953.9712897
- Sternberg, R. J., and Lubart, T. I. (1999). “The concept of creativity: prospects and paradigms” in *Handbook of creativity*. ed. R. J. Sternberg (New York, NY: Cambridge University Press), 3–15.
- Stevenson, C., Smal, I., Baas, M., Dahrendorf, M., Grasman, R., Tanis, C., et al. (2020). Automated AUT scoring using a big data variant of the consensual assessment technique: Final technical report. Amsterdam: Universiteit van Amsterdam.
- Stevenson, C., Smal, I., Baas, M., Grasman, R., and van der Maas, H. (2023). Putting GPT-3’s creativity to the (alternative uses) test. Available online at <https://arxiv.org/abs/2206.08932> (Accessed June 15, 2025).
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personal. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1
- Sun, J., Chen, Q., Zhang, Q., Li, Y., Li, H., Wei, D., et al. (2016). Training your brain to be more creative: brain functional and structural changes induced by divergent thinking training. *Hum. Brain Mapp.* 37, 3375–3387. doi: 10.1002/hbm.23246
- Sun, L., Gu, H., Myers, R., and Yuan, Z. (2024). “A new dataset and method for creativity assessment using the alternate uses task” in *Intelligent computers, algorithms, and applications: Third bench council international symposium, IC 2023*. eds. C. Cruz, Y. Zhang and W. Gao (Singapore: Springer), 125–138.
- Sundar, S. S. (2008). “The MAIN model: a heuristic approach to understanding technology effects on credibility” in *Digital media, youth, and credibility*. eds. M. J. Metzger and A. J. Flanagin (Cambridge, MA: MIT Press), 73–100.
- Tang, S., Qian, W., Liu, P., and Cao, J. (2025). Art creator: steering styles in diffusion model. *Neurocomputing* 626:129511. doi: 10.1016/j.neucom.2025.129511
- Threadgold, E., Marsh, J. E., and Ball, L. J. (2018). Normative data for 84 UK English rebus puzzles. *Front. Psychol.* 9:2513. doi: 10.3389/fpsyg.2018.02513
- Threadgold, E., Marsh, J. E., McLatchie, N., and Ball, L. J. (2019). Background music stints creativity: evidence from compound remote associate tasks. *Appl. Cognit. Psych.* 33, 873–888. doi: 10.1002/acp.3532
- Tian, Y., Ravichander, A., Qin, L., Bras, R. L., Marjeh, R., Peng, N., et al. (2023). Macgyver: are large language models creative problem solvers?. Available online at: <https://arxiv.org/pdf/2311.09682> (Accessed June 15, 2025).
- Tigre Moura, F., Castrucci, C., and Hindley, C. (2023). Artificial intelligence creates art? An experimental investigation of value and creativity perceptions. *J. Creat. Behav.* 57, 534–549. doi: 10.1002/jocb.600
- Toh, C. A., and Miller, S. R. (2016). Choosing creativity: the role of individual risk and ambiguity aversion on creative concept selection in engineering design. *Res. Eng. Des.* 27, 195–219. doi: 10.1007/s00163-015-0212-1
- Toivainen, T., Olteteanu, A.-M., Repeykova, V., Lihanov, M., and Kovas, Y. (2019). Visual and linguistic stimuli in the remote associates test: a cross-cultural investigation. *Front. Psychol.* 10:926. doi: 10.3389/fpsyg.2019.00926
- Tresset, P., and Leymarie, F. F. (2013). Portrait drawing by Paul the robot. *Comput. Graph.* 37, 348–363. doi: 10.1016/j.cag.2013.01.012
- Tromp, C., and Glăveanu, V. P. (2023). A model of creative aging (MOCA): unlocking the potential of constraints for creativity in older adults. *Psychol. Aesthet. Creat. Arts* 19, 247–256. doi: 10.1037/aca0000538
- Umiltà, M. A., Berchio, C., Sestito, M., Freedberg, D., and Gallese, V. (2012). Abstract art and cortical motor activation: an EEG study. *Front. Hum. Neurosci.* 6:311. doi: 10.3389/fnhum.2012.00311
- Urban, K. K. (2005). Assessing creativity: the test for creative thinking- drawing production (TCT-DP). *Int. Educ. J.* 6, 272–280.
- Urban, K. K., and Jellen, H. G. (1996). Test for creative thinking-drawing production (TCT-DP). Lisse, Netherlands: Swets & Zeitlinger.
- van Esch, P., Black, J. S., and Ferolie, J. (2019). Marketing AI recruitment: the next phase in job application and selection. *Comput. Hum. Behav.* 90, 215–222. doi: 10.1016/j.chb.2018.09.009
- Van Geert, E., and Wagemans, J. (2020). Order, complexity, and aesthetic appreciation. *Psychol. Aesthet. Creat. Arts* 14, 135–154. doi: 10.1037/aca0000224
- van Welzen, K., Ullén, F., and de Manzano, Ö. (2024). Does divergent thinking relate to expertise? Introducing a novel test of creative ideation in music. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000715
- Velásquez-Salamanca, D., Martín-Pascual, M. Á., and Andreu-Sánchez, C. (2025). Interpretation of AI-generated vs. human-made images. *J. Imaging* 11:227. doi: 10.3390/jimaging11070227
- von Krogh, G. (2018). Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing. *Acad. Manag. Discov.* 4, 404–409. doi: 10.5465/amd.2018.0084
- Wallach, M. A., and Kogan, N. (1965). Modes of thinking in young children. New York: Holt, Rinehart and Winston.
- Waytz, A., Heafner, J., and Epley, N. (2014). The mind in the machine: anthropomorphism increases trust in an autonomous vehicle. *J. Exp. Soc. Psychol.* 52, 113–117. doi: 10.1016/j.jesp.2014.01.005
- Webb, M. E., Little, D. R., and Cropper, S. J. (2016). Insight is not in the problem: investigating insight in problem solving across task types. *Front. Psychol.* 7:1424. doi: 10.3389/fpsyg.2016.01424
- Webb, M. E., Little, D. R., and Cropper, S. J. (2021). Unusual uses and experiences are good for feeling insightful, but not for problem solving: contributions of schizotypy, divergent thinking, and fluid reasoning, to insight moments. *J. Cogn. Psychol.* 33, 770–792. doi: 10.1080/20445911.2021.1929254
- Weiss, S., and Wilhelm, O. (2022). Is flexibility more than fluency and originality? *J. Intelligence* 10:96. doi: 10.3390/jintelligence10040096
- Williams, F. E. Creativity assessment packet: CAP Pro-Ed. (1993).

- Winkielman, P., and Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect. *J. Pers. Soc. Psychol.* 81, 989–1000. doi: 10.1037/0022-3514.81.6.989
- Winkielman, P., Halberstadt, J., Fazendeiro, T., and Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychol. Sci.* 17, 799–806. doi: 10.1111/j.1467-9280.2006.01785.x
- Wu, C. L., Huang, S. Y., Chen, P. Z., and Chen, H. C. (2020). A systematic review of creativity-related studies applying the remote associates test from 2000 to 2019. *Front. Psychol.* 11:573432. doi: 10.3389/fpsyg.2020.573432
- Wu, Y., Mou, Y., Li, Z., and Xu, K. (2020). Investigating American and Chinese subjects' explicit and implicit perceptions of AI-generated artistic work. *Comput. Hum. Behav.* 104:106186. doi: 10.1016/j.chb.2019.106186
- Wurtz, P., Reber, R., and Zimmermann, T. D. (2008). The feeling of fluent perception: a single experience from multiple asynchronous sources. *Conscious. Cogn.* 17, 171–184. doi: 10.1016/j.concog.2007.07.001
- Xing, Q., Lu, Z., and Hu, J. (2019). The effect of working memory updating ability on spatial insight problem solving: evidence from behavior and eye movement studies. *Front. Psychol.* 10:927. doi: 10.3389/fpsyg.2019.00927
- Xing, Q., Rong, C., Lu, Z., Yao, Y., Zhang, Z., and Zhao, X. (2018). The effect of the embodied guidance in the insight problem solving: an eye movement study. *Front. Psychol.* 9:2257. doi: 10.3389/fpsyg.2018.02257
- Xue, A. (2021). End-to-end Chinese landscape painting creation using generative adversarial networks. In *Proceedings of the IEEE/CVF Winter conference on applications of computer vision*. Available online at: https://openaccess.thecvf.com/content/WACV2021/papers/Xue_End-to-End_Chinese_Landscape_Painting_Creation_Using_Generative_Adversarial_Networks_WACV_2021_paper.pdf (Accessed June 15, 2025).
- Yang, Y., and Xu, H. (2025). Perception of AI creativity: dimensional exploration and scale development. *J. Creat. Behav.* 59:e70028. doi: 10.1002/jocb.70028
- Yeomans, M., Shah, A., Mullainathan, S., and Kleinberg, J. (2019). Making sense of recommendations. *J. Behav. Dec. Making* 32, 403–414. doi: 10.1002/bdm.2118
- Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *Am. Psychol.* 35, 151–175. doi: 10.1037/0003-066X.35.2.151
- Zamani, A., Mills, C., Girn, M., and Christoff, K. (2023). “A closer look at transitions between the generative and evaluative phases of creative thought” in *The Routledge international handbook of creative cognition*. eds. L. J. Ball and F. Vallée-Tourangeau (London, UK: Routledge), 453–474.
- Zedelius, C. M., and Schooler, J. W. (2015). Mind wandering “Ahas” versus mindful reasoning: alternative routes to creative solutions. *Front. Psychol.* 6:834. doi: 10.3389/fpsyg.2015.00834
- Zhou, J., Wang, X. M., Song, L. J., and Wu, J. (2017). Is it new? Personal and contextual influences on perceptions of novelty and creativity. *J. Appl. Psychol.* 102, 180–202. doi: 10.1037/apl0000166