Evaluating Creativity: How Idea Context and Rater Personality Affect Considerations of Novelty and Usefulness

James Lloyd-Cox, Alan Pickering & Joydeep Bhattacharya

Published online: 30 Sep 2022.

Submit your article to this journal ↗

Article views: 75

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Evaluating Creativity: How Idea Context and Rater Personality Affect Considerations of Novelty and Usefulness

James Lloyd-Cox ⓘD, Alan Pickering ⓘD, and Joydeep Bhattacharya ⓘD

Goldsmiths, University of London

ABSTRACT

According to the standard definition, creative ideas must be both novel and useful. While a handful of recent studies suggest that novelty is more important than usefulness to evaluations of creativity, little is known about the contextual and interpersonal factors that affect how people weigh these two components when making an overall creativity judgment. We used individual participant regressions and mixed-effects modeling to examine how the contributions of novelty and usefulness to ratings of creativity vary according to the context of the idea (i.e., how relevant it is to the real world) and the personality of the rater. Participants ($N = 121$) rated the novelty, usefulness, and creativity of ideas from two contexts: responses to the alternative uses task (AUT) and genuine suggestions for urban planning projects. We also assessed three personality traits of participants: openness, intellect, and risk-taking. We found that novelty contributed more to evaluations of creativity among AUT ideas than projects, while usefulness contributed more among projects than AUT ideas. Further, participants with higher openness and higher intellect placed a greater emphasis on novelty when evaluating AUT ideas, but a greater emphasis on usefulness when evaluating projects. No significant effects were found for the risk-taking trait.

PLAIN LANGUAGE SUMMARY

Understanding how creativity is perceived and defined in different contexts and across different individuals is highly important, not just to our understanding of how to assess creativity, but to our understanding of creativity itself. However, relatively few existing studies have examined differences in how individuals evaluate creativity, and the factors they consider during their evaluations. We investigated how personality and problem context affect how individuals consider novelty and usefulness when making an overall creativity judgment. Participants rated ideas from two contexts: responses to a common lab-based measure of creative ability and genuine suggestions for urban planning projects (ideas with more real-world relevance). We also assessed three personality traits of participants: openness, intellect, and risk-taking. Data was analyzed using both individual participant regressions and linear mixed-effects models. We found that participants considered novelty more when evaluating the creativity of AUT ideas (relative to projects), and usefulness more when evaluating the creativity of projects (relative to AUT ideas). Furthermore, when evaluating the creativity of AUT ideas, participants with higher openness and higher intellect placed a greater emphasis on novelty, but when evaluating projects, these same participants placed a greater emphasis on usefulness. Overall, our findings highlight the importance of considering contextual and interpersonal factors when researchers examine how creativity is evaluated, defined, and perceived, strengthening recent calls for creativity assessments that can account for variation across raters.

Introduction

Creativity is a hugely important, yet somewhat mysterious ability that enables humans to craft innovative solutions, adopt original perspectives, to invent, imagine and entertain. However, defining precisely what creativity is and what constitutes a creative idea has proven difficult and controversial (Plucker, Beghetto, & Dow, 2004; Simonton, 2018; Taylor, 1988; Treffinger, 1992), with considerable variation in the working definition of

creativity across fields of research (Hennessey & Amabile, 2010; Puryear & Lamb, 2020).

Perhaps the most commonly accepted definition of creativity is the "standard definition" (Runco & Jaeger, 2012), which states that to be creative, an idea must be both novel and useful. Though the precise terminology can vary (e.g., novelty may be referred to as originality or uniqueness, while usefulness may be referred to as appropriateness, relevance, or effectiveness), the twin

CONTACT James Lloyd-Cox ✉ jlloy001@gold.ac.uk 🖳 Department of Psychology, Goldsmiths, University of London, New Cross, London, SE14 6NW, UK

criteria of novelty and usefulness have formed principal components of numerous definitions of creativity dating back at least 70 years (Amabile, 1982; Plucker et al., 2004; Stein, 1953). The definition is not without conceptual issues (see Corazza, 2016; Martin & Wilson, 2017), and some have suggested additional requirements including surprise (Boden, 2007; Simonton, 2018), discovery (Martin & Wilson, 2017), and aesthetics and authenticity (Kharkhurin, 2014). However, especially within cognitive psychology and neuroscience, the standard definition continues to provide a theoretical foundation for vast amounts of creativity research, and to serve as a guide when raters evaluate the creativity of ideas, products, or responses.

If a creative idea is (at minimum) both novel and useful, it seems likely that when evaluating the creativity of an idea, raters would make their final judgment based on a certain weighting of its perceived novelty and usefulness. However, surprisingly little research has investigated how these components contribute to evaluations of creativity, and the factors that can modify these contributions. While some research suggests that novelty is far more important to creativity than usefulness (Caroff & Besançon, 2008; Diedrich, Benedek, Jauk, & Neubauer, 2015; Han, Forbes, & Schaefer, 2021; Runco & Charles, 1993), other findings indicate that the contributions of novelty and usefulness may depend on the context in which the idea was generated and the nature of the problem it is intended to solve (Acar, Burnett, & Cabra, 2017; Long, 2014; Runco, Illies, & Eisenman, 2005).

Meanwhile, although researchers have examined how individual differences, including expertise (Long, 2014), emotion (Lee, Chang, & Choi, 2017; Mastria, Agnoli, Corazza, & Eisenbarth, 2019), and uncertainty (Mueller, Melwani, & Goncalo, 2012), can influence evaluations of overall creativity, little is known about how these differences might affect considerations of novelty and usefulness. Personality, particularly the Big-Five trait openness/intellect, is likely to be an important factor here, since it determines how receptive individuals are to new and unusual ideas (Kaufman et al., 2016; Oleynick et al., 2017), potentially driving them to consider novelty more than usefulness when they evaluate creativity. However, it remains unknown how factors such as the nature of the creative task and the personality of the rater can affect how novelty and usefulness contribute to evaluations of creativity. Providing answers to these questions is of central importance to our understanding of how creativity is evaluated, defined, and perceived, and may inform the development of subjective creativity assessments that can account for variance across raters (Barbot, Hass, &

Reiter-Palmon, 2019; Myszkowski & Storme, 2019). As a brief but important note, this study is concerned with the evaluation of exogenous ideas (i.e., ideas generated by others) as opposed to the evaluation of one's own ideas, which is likely to be a related but distinct evaluative process (Karwowski, Czerwonka, & Kaufman, 2020; Rodriguez, Cheban, Shah, & Watts, 2020; Runco & Smith, 1992).

Assessing creativity and its components

While creativity can be assessed through self-report methods that focus on creative achievements and hobbies (e.g., Carson, Peterson, & Higgins, 2005; Diedrich et al., 2018; Kaufman, 2019), lab-based creativity tests typically require participants to produce creative responses or products, such as musical improvisations (Pinho, de Manzano, Fransson, Eriksson, & Ullén, 2014), drawings (Rominger et al., 2018), or short stories (Prabhakaran, Green, & Gray, 2014), which are then evaluated by a panel of raters. One of the most common creativity tests is the alternative uses task (AUT), which requires participants to think of unusual uses for everyday objects, such as a brick or a table (Gilhooly, Fioratou, Anthony, & Wynn, 2007; Guilford, 1967). When it comes to evaluating creativity as a single, holistic construct, the gold-standard method within psychology is the consensual assessment technique (CAT; Amabile, 1982; Baer & McKool, 2014; Kaufman, Lee, Baer, & Lee, 2007; see also Cseh & Jeffries, 2019), in which several expert judges rate the creativity of each idea on a Likert scale. Ratings are then averaged across raters.

As mentioned earlier, creativity has two essential components – novelty and usefulness. Novelty refers to the unusualness, uniqueness, and originality of an idea and can be assessed either through subjective ratings (e.g., Acar et al., 2017; Diedrich et al., 2015; Silvia, 2008) or through objective measures such as the statistical infrequency of the idea among the current sample (Plucker, Qian, & Wang, 2011; Runco et al., 2005; Wilson, Guilford, & Christensen, 1953). By contrast, usefulness refers to the feasibility, appropriateness, and value of an idea, which in the majority of tasks can be determined only by subjective assessment (Acar et al., 2017; Diedrich et al., 2015; Runco et al., 2005).

How novelty and usefulness contribute to evaluations of creativity: the role of idea context

How do novelty and usefulness contribute to evaluations of creativity, and is one component more important than the other? Novelty and usefulness ratings are

often negatively correlated (Caroff & Besançon, 2008; Diedrich et al., 2015; Runco & Charles, 1993), so an optimally creative idea may have to balance a trade-off between novelty and usefulness. To date, however, only a handful of studies have examined how novelty and usefulness contribute to evaluations of creativity. The majority of this research has found that the perceived creativity of an idea depends more on its novelty that on its usefulness, in contexts including AUT ideas (Acar et al., 2017; Diedrich et al., 2015; Runco & Charles, 1993), advertisements (Caroff & Besançon, 2008; Storme & Lubart, 2012), and product designs (Han et al., 2021). For example, Diedrich et al. (2015) asked 18 participants to rate the novelty, usefulness, and creativity of around 5000 ideas produced in both the AUT and a figural-completion drawing task. They found that creativity ratings were far more strongly related to novelty ratings (with β estimates ranging between .75 and .81) than usefulness ratings (β estimates between .26 and .32). They also found a significant interaction between novelty and usefulness, whereby usefulness was less related to creativity among common (i.e., non-novel) ideas and far more related to creativity among novel ideas.

However, some findings suggest that the contributions of novelty and usefulness to evaluations of creativity may depend on the context in which the idea is produced. Runco and colleagues (2005) examined ideas for both realistic problems (with potential application to the real-world) and unrealistic problems (unlikely to be encountered in the real world). Ideas for realistic problems were rated as more useful than ideas for unrealistic problems, while ideas for unrealistic problems were rated as more novel. While relations with creativity were not examined, these findings indicate that certain contexts may elicit a different consideration of novelty and usefulness when raters evaluate creativity. For example, usefulness may have a minimal impact on evaluations of creativity in contexts where it is less relevant (such as with adverts, artworks, and AUT ideas) but may draw far more consideration in the context of genuine real-world problems. This possibility is further supported by a qualitative study, which found that when raters evaluated the creativity of scientific ideas, novelty and usefulness were considered equally important criteria (Long, 2014).

A further suggestion that the relationships between novelty, usefulness, and creativity might depend on the context of the creative idea comes from Acar et al. (2017), who examined how four factors, including novelty and usefulness, contributed to judgments of creativity. In their study, 776 participants completed ratings for both AUT ideas and real-world creative products. The authors again found novelty to be more related to creativity than usefulness, but also found evidence that the relationship between usefulness and creativity may depend on the context of the idea. However, the study focused on variance at the rater level, examining ratings for only 12 ideas (all of which had high prior ratings of creativity), and the results were inconclusive as to which context displayed the greater relationship between usefulness and creativity. To our knowledge, no study has provided definitive evidence regarding how the context of ideas can affect the contributions of novelty and usefulness to evaluations of creativity.

Individual differences in the evaluation of creativity and its components

In addition to the context of the idea, individual differences between raters are also likely to influence the contributions of novelty and usefulness to evaluations of creativity. Understanding differences in the evaluation of creativity is highly important to creativity research for at least two reasons. First, creativity research relies heavily on subjective assessments of creativity, and so understanding the interpersonal factors that cause variation in these assessments is key to developing strong and reliable measures. Indeed, the most common subjective assessment method, the CAT, has recently been criticized for not accounting for variation across raters (Barbot et al., 2019; Myszkowski & Storme, 2019). The limitations intrinsic to subjective assessments of creativity are well-known, and have stimulated the development of objective assessments including distributional semantics methods (Acar, Berthiaume, Grajzel, Dumas, & Flemister, 2021; Beaty & Johnson, 2021) and machine learning techniques (Cropley & Marrone, 2021; Edwards, Peng, Miller, & Ahmed, 2021). However, such methods can often assess only the novelty of ideas, not the usefulness (Beaty & Johnson, 2021), and the field will likely continue to rely on subjective assessments of creativity for the foreseeable future.

Second, a better understanding of creative evaluation could lead to a better understanding of creative generation. The production of creative ideas is often argued to involve iterative cycles of generation and evaluation (e.g., Basadur, 1995; Finke, Ward, & Smith, 1992; Lubart, 2001; cf. Campbell, 1960; Simonton, 2013), and research suggests that more thorough evaluation during the production of ideas can lead to better creative performance (Gibson & Mumford, 2013; McIntosh, Mulhearn, & Mumford, 2021; Watts, Steele, Medeiros, & Mumford, 2019). Moreover, given the close ties between generation and evaluation, differences in how

people evaluate ideas may relate to differences in how people generate ideas. For example, individuals who favor novelty over usefulness when evaluating the ideas of others may show the same preferences when generating their own products or responses (a possibility supported by Caroff & Besançon, 2008). As such, a clearer understanding of differences in the evaluation of creativity may lead not only to more nuanced creative assessment techniques but also to a clearer understanding of differences in creative performance.

A considerable body of work has examined how differences in the evaluation of creative ideas relate to factors including culture (Ivancovsky, Shamay-Tsoory, Lee, Morio, & Kurman, 2019; McCarthy, Chen, & McNamee, 2018; Simonton, 1999; Sternberg, 2018), intelligence (Karwowski et al., 2020; Storme & Lubart, 2012), musical training (Kleinmintz et al., 2014; Kleinmintz, Ivancovsky, & Shamay-Tsoory, 2019), emotion (Mastria et al., 2019), and uncertainty (Lee et al., 2017; Mueller et al., 2012). Of particular note, research suggests that positive emotion may relate to higher creativity ratings (Mastria et al., 2019), while uncertainty relates to lower creativity ratings (Lee et al., 2017; Mueller et al., 2012). It has also been found that prevention focus (a tendency to minimize loss) is related to greater accuracy when evaluating usefulness and reduced accuracy when evaluating novelty, compared to promotion focus (a tendency to maximize reward; Herman & Reiter-Palmon, 2011). These findings suggest that more negative, uncertain, and avoidance-oriented states may lead raters to favor practicality over creativity, shunning novel ideas that may be associated with greater risk. By contrast, more positive, certain, and promotion-oriented states might lead raters to be more receptive to creative and novel ideas. In line with this research, it seems likely that an individual's personality and preference for risk-taking might also impact how they evaluate creativity, and indeed how they weigh novelty and usefulness when evaluating creativity.

Research into the link between personality and creativity has a rich history (Batey & Furnham, 2006; Feist, 1998). In particular, the Big Five trait openness/intellect has been found to relate to greater scores on virtually all forms of creativity assessment (Batey & Furnham, 2006; Feist, 1998; Kaufman et al., 2016; Oleynick et al., 2017). Openness/intellect is typified by imagination and artistic and intellectual curiosity, and may be assessed as a single construct or in terms of its twin aspects of openness and intellect (Kaufman et al., 2016; Oleynick et al., 2017). Among possible reasons for the link between greater openness/intellect and greater creativity is that those higher in the trait tend to seek out novelty and complexity, and are motivated by a recurrent desire to enlarge

their experience (DeYoung, Peterson, & Higgins, 2005; Kaufman et al., 2016; McCrae & Ingraham, 1987; Oleynick et al., 2017).

Given this characterization, it seems possible that individuals with higher openness/intellect scores might be more receptive to creative ideas, and may place more importance on novelty and less on usefulness when evaluating creativity. It is also possible that openness and intellect, examined separately, are associated with different weightings of novelty and usefulness. For example, Kaufman et al. (2016) found that while openness predicts creative achievement in the arts, intellect predicts creative achievement in the sciences. As such, one might expect openness to relate to a greater consideration of novelty and intellect to relate to a greater consideration of usefulness when participants evaluate creativity. However, while some research has investigated how openness/intellect relates to the evaluation of creativity overall (Ceh, Edelmann, Hofer, & Benedek, 2022; Silvia, 2008), very little is known about how differences in openness/intellect relate to differences in how novelty and usefulness are weighted during the evaluation of creativity.

An individual's willingness to take risks might also affect how they evaluate creative ideas. By definition, creative ideas are different from the norm, and as their novelty increases, they may be less likely to be appropriate or useful (as is indicated by the negative relationship often found between novelty and usefulness; Caroff & Besançon, 2008; Diedrich et al., 2015; Runco & Charles, 1993). As such, individuals who are more willing to take risks might be more willing to pursue creative ideas, and may place more weight on novelty than usefulness when assessing creativity. The relationship between risk-taking and creativity is not as clear-cut as for openness/intellect, with some studies finding a positive relationship (Dewett, 2007; Glover & Sautter, 1977) and others finding no relationship (Erbas & Bas, 2015; Shen, Hommel, Yuan, Chang, & Zhang, 2018). More recent research suggests that it may be social risk-taking (and not risk-taking in other domains) that relates to greater creativity (Bonetto, Pichot, Pavani, & Adam-Troïan, 2021; Tyagi, Hanoch, Hall, Runco, & Denham, 2017). However, it remains unknown how risk-taking affects the evaluation of creative ideas and the importance assigned to novelty and usefulness.

The present research

Empirical and theoretical work suggests that creative ideas are both novel and useful. However, while some research indicates that novelty is more important than usefulness to evaluations of creativity, it remains

unknown how the contributions of these components depend on the nature of the creative task and how applicable it is to the real world. In addition, despite the importance of subjective assessments to creativity research, it is unclear how individual differences among raters can affect their evaluations of creativity and the importance they assign to novelty and usefulness.

To investigate these outstanding questions, we followed a hierarchical, mixed-effects design (with ratings nested within participants) to examine how idea context and rater personality can affect the contributions of novelty and usefulness to evaluations of creativity. Participants rated the novelty, usefulness, and creativity of ideas from two contexts: AUT ideas and genuine suggestions for social development projects (subsequently referred to as "Projects"). Following these ratings, participants completed questionnaires assessing openness/intellect and risk-taking traits. Relationships between idea ratings and personality scores were then examined using both single-subject maximum likelihood estimation (SSMLE) and linear mixed-effects models (LMEMs)

We had several predictions, in line with the hypothesis that when evaluating creativity, raters would weigh the novelty and usefulness of an idea differently depending on their personality traits and the idea's context (i.e., real-world relevance). Concerning context, we predicted that among AUT ideas, creativity ratings would be more related to novelty ratings than usefulness ratings (as found previously; Diedrich et al., 2015; Runco & Charles, 1993; Storme & Lubart, 2012). However, consistent with the notion that usefulness is a more important component of creativity in the context of more realistic problems (Long, 2014; Runco et al., 2005), we also predicted that creativity ratings would be more related to usefulness ratings among Projects than among AUT ideas.

Concerning personality traits, we predicted that openness and risk-taking would both be associated with a stronger relationship between novelty and creativity, among idea ratings in both contexts. This would be consistent with the notion that individuals who are more open to new ideas, and more likely to take risks, are more driven toward novelty and so value novelty more when evaluating creativity. We had no specific predictions regarding how intellect would moderate relationships; however, we wished to examine whether higher intellect scores would be associated with a stronger relationship between usefulness and creativity, given research linking intellect to creative achievement in the sciences, but not in the arts (Kaufman et al., 2016).

Methods

Participants

Using G*power software (version 3.1; Faul, Erdfelder, Lang, & Buchner, 2007), we calculated that a sample size of 111 was required for a 95% power to detect correlations of $r = .03$ or greater. As such, 121 healthy human adults (88 females; mean age = 31.3 years, $SD = 14.3$) were recruited for the study. 36 were recruited from Goldsmiths, University of London and did not receive any financial incentive, while 85 were recruited via Prolific and were paid a small cash incentive. Among paid participants, participation was contingent on a Prolific approval rating of 90% or above and a minimum of 40 previously completed studies. Fluency in English was required for participants in both samples due to the nature of the task, which involved evaluating the creativity of verbal ideas. Among both paid and non-paid participants, informed consent was given prior to data collection. Ethical approval for the study was given by the Local Ethics Committee of the Department of Psychology at Goldsmiths, University of London.

Materials

Idea ratings: AUT responses

AUT ideas were 48 suggested uses for one of two objects: "table" and "shoe." The ideas were carefully selected from a total of 1866 responses produced by participants in a prior study (Luft, Zioga, Thompson, Banissy, & Bhattacharya, 2018), to ensure an even distribution in terms of creative quality. Each idea had been rated for creativity on a scale from 1 to 10 by three raters. For the present study, scores were averaged across these raters to produce one creativity score per idea. Ideas were then spelling-corrected, and repeated items were removed. Next, histograms of idea creativity were examined for each object. Ideas for both objects were highly skewed, with very few ideas scoring above 8 in creativity. To produce more even distributions, ratings of 9 or 10 were recoded as 8. Ideas were then separated into four bins, each corresponding to a rating of 1–2, 3–4, 5–6, and 7–8. For each object, 48 ideas were pseudorandomly selected (for a total of 96), such that 12 ideas came from each rating bin. These were then manually checked, and inappropriate or very similar ideas were removed, leaving 24 ideas per object (48 in total). Finally, ideas were rephrased for succinctness.

Idea ratings: social development projects

Projects were 10 suggestions for urban planning projects that might "restore vibrancy in cities and regions facing

economic decline." During a prior study (Pétervári, 2018), Projects had been selected from an open-source platform (OpenIDEO, 2011) from among entrees into a competition, and reduced to two-paragraph descriptions. Participants in this prior study ($N = 80$) rated their willingness to invest in each Project on a scale from 0 to 100, which was assumed to indicate the Project's overall quality. For the present study, 10 Projects were selected from a total of 15, due to time constraints and the longer length of the Project descriptions compared to the AUT ideas. This was achieved by removing the 5 Projects with the most variable ratings, which increased the uniformity of the quality scores across Projects.

Openness/intellect

The Openness/Intellect subscale of the Big Five Aspect Scale (BFAS; DeYoung, Quilty, & Peterson, 2007) was used to assess openness and intellect. The subscale contains 20 items, 10 of which assess openness and 10 of which assess intellect, allowing the two facets to be examined separately. Each item is a statement (e.g., "I am quick to understand things"). Participants indicate their agreement with each statement using a 1 (strongly disagree) to 5 (strongly agree) scale.

Risk-taking

The Domain Specific Risk-taking Scale (DSRS; Blais & Weber, 2006) was used to assess risk-taking. The scale comprises 30 items, with six items for each of five domains of risk-taking: ethical, financial, health/safety, recreational, and social. Items are descriptions of activities or behaviors (e.g., "bungee jumping off a tall bridge"), and participants must indicate how likely they would be to engage in the activity using a scale from 1 (extremely unlikely) to 7 (extremely likely). Participant scores are summed across the five domains to produce a single, general risk-taking score. However, in line with research that suggests it is social risk-taking specifically that relates to creativity (Bonetto et al., 2021; Tyagi et al., 2017), we included participants' scores on both general risk-taking and social risk-taking in our analysis.

Procedure

All data was collected using Qualtrics software. Participants completed idea ratings first, and personality measures second. Idea rating trials were organized into blocks by idea context (i.e., AUT or Projects), and by property (i.e., novelty, usefulness, or creativity). Participants completed blocks in one of four orders to counterbalance the order of contexts and properties. Specifically, half of the participants completed AUT ratings first, while half completed Projects ratings first. Within these groups, half of the participants completed novelty ratings first, while the other half completed usefulness ratings first. All participants completed ratings for overall creativity last, though the order of AUT ideas and Projects varied within creativity ratings (see Figure 1). Within each block (e.g., novelty ratings for AUT ideas) trials were randomized.

Participants were initially told only that they would be "evaluating ideas." No instructions regarding novelty, usefulness, or creativity were given until participants began the corresponding block. As such, participants were naive to the fact that they would be rating creativity until after they had completed both novelty and usefulness ratings. Upon starting each block, participants were told they would be asked to rate the novelty, usefulness, or creativity of either "ideas for how to use common, everyday objects" or "real proposals for urban planning projects." Participants were then given further instructions to help them consider the property in question. Specifically, for novelty, usefulness, and creativity respectively, they were told to think about: "how novel, unusual, or unexpected each idea is"; "how useful, effective, or practical each idea is"; or "how creative each idea is." Since instructions pertaining to creativity often ask participants to focus on originality, novelty, usefulness, or appropriateness (Acar, Runco, & Park, 2019), and since these components were being rated separately in our study, creativity was deliberately left open to interpretation, with minimal additional instructions. Participants then completed two (Projects) or five (AUT) practice ratings before seeing the same instructions again and beginning the real trials. Instructions
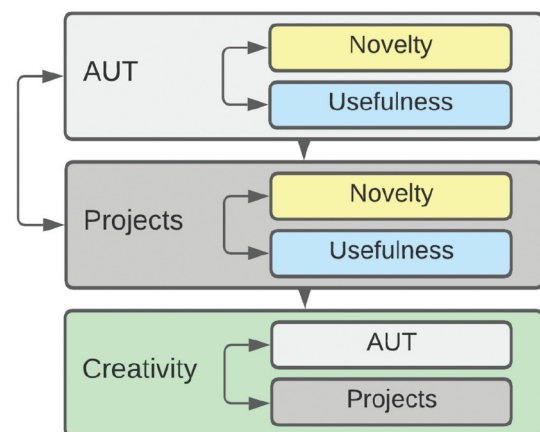

Figure 1. Order of rating blocks (top to bottom). The order of contexts (i.e., AUT ideas or Projects) and properties (i.e., novelty or usefulness) were counterbalanced. Double-ended arrows denote interchangeability dependent on counterbalancing conditions. Creativity ratings were always completed last.

were repeated to emphasize the points participants should consider in their ratings.

Within each trial, participants were shown a single line of instruction (e.g., "How NOVEL is this idea for how to restore vibrancy in cities and regions facing economic decline?," or "How USEFUL is this idea for how to use a table?"), together with the idea itself, and a scale from 1 (not at all) to 7 (very). After finishing all ratings, participants completed questionnaires assessing openness/intellect and risk-taking.

Analyses

Analyses made use of both SSMLE and LMEMs. LMEMs can account for the dependence of multiple data points from a single individual (here, ratings for different ideas), modeling them as random effects (Singmann & Kellen, 2019). This allowed us to model unique relationships between novelty, usefulness, and creativity for each participant, while simultaneously estimating group-level effects (McNeish & Kelley, 2019). By contrast, SSMLE (Katahira, 2016) is a more intuitive approach that involves fitting a standard linear regression for each participant separately. While this approach is known to be generally less powerful than LMEMs (see Stein's paradox; Efron & Morris, 1977; Katahira, 2016), SSMLE provides distributions of predictor estimates (e.g., for novelty and usefulness) which can then be compared for significant differences, while correlations can be computed between parameter estimates and individual differences (e.g., openness). The two forms of analysis have different assumptions, and so using both together can provide a richer understanding of the examined relationships, as well as an indication of the robustness of findings (see Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016).

SSMLE was conducted first, and separately for AUT ideas and Projects, to compare the relative importance of novelty and usefulness when evaluating creativity in both contexts. Regressions were fitted for each participant individually, with creativity as the dependent variable and novelty and usefulness as joint predictors (i.e., novelty and usefulness were simultaneously present in each regression). Prior to computing regressions, creativity, novelty, and usefulness ratings were z-scored within participants. The standardized beta coefficients for novelty and usefulness were then used in further analyses, to compare the coefficients between idea contexts, and to examine relationships between coefficients and personality measures.

Following the SSMLE analyses, a series of LMEMs were computed to further examine the relationships between creativity, novelty, and usefulness ratings, and

to test whether these relationships were significantly moderated by context and participants' personality scores. In addition, we wished to test for a significant interaction between novelty and usefulness, as has been found previously (Diedrich et al., 2015). Three LMEMs were computed, all of which had creativity rating as the dependent variable. These models were constructed by successively adding effects to create more complex versions of the model, comparing each model to the previous, simpler model via likelihood ratio testing (e.g., Wilken, Forthmann, & Holling, 2020). This should reveal whether each added effect contributes significantly to model fit. Models were computed using custom MATLAB scripts and the fitlme function. As with SSMLE, creativity, novelty, and usefulness were z-scored within participants. In addition, personality scores were z-scored across participants.

Results

Of 121 participants, data for 9 were removed due to these participants failing attention checks or responding randomly. One additional participant's data was removed from all analyses involving the Projects ratings due to incomplete data for this part of the study. The final sample sizes were thus 112 for the AUT data and 111 for the Projects data.

To check for differences between paid ($N = 76$) and non-paid ($N = 36$) samples, a series of independent samples t-tests were conducted. No significant differences were found between paid and non-paid participants, either in terms of novelty, usefulness, or creativity ratings (among either AUT ideas or Projects) or in terms of personality measures ($p > .235$ in all cases).

Descriptive statistics

Descriptive statistics and zero-order correlations for participants' personality scores and novelty, usefulness, and creativity ratings are shown in Table 1. Here, idea ratings are averaged within participants to produce a single score for each rating block and each participant.

Creativity was positively related to novelty, among both AUT ($r = .51$, $p < .001$) and Project ratings ($r = .60$, $p < .001$). By contrast, creativity was positively related to usefulness only among Project ratings ($r = .61$, $p < .001$), not AUT ratings ($r = .03$, $p = .758$). In addition, novelty and usefulness ratings were positively correlated only among Project ratings ($r = .49$, $p < .001$). While correlations between personality measures and participants' mean ratings were of secondary interest in this study (which is primarily interested in how personality measures moderate relationships between creativity

Table 1. Means, standard deviations, and correlation coefficients for personality measures and participant-level idea ratings.

	Mean	SD	1	2	3	4	5	6	7	8	9
1. Openness	38.63	5.47	-								
2. Intellect	37.11	5.79	.32**	-							
3. Risk General	93.88	22.33	−.04	.18	-						
4. Risk Social	30.81	5.21	.19*	.31**	.47**	-					
5. AUT Nov.	4.03	0.60	.02	.12	.00	−.10	-				
6. AUT Use.	3.81	0.57	−.03	−.14	−.04	.06	−.09	-			
7. AUT Crea.	3.94	0.66	.10	.13	.01	−.13	.51**	.03	-		
8. Proj. Nov.	4.57	0.81	.03	.03	−.11	.13	.25**	.18	.31**	-	
9. Proj. Use.	4.67	0.75	.02	−.06	.07	.14	.05	.31**	.15	.49**	-
10. Proj. Crea.	4.88	0.79	.15	−.07	.05	.12	.10	.13	.31**	.60**	.61**

Note. Risk General = general risk-taking; Risk Social = social risk-taking; Nov. = novelty; Use. = usefulness; Crea. = creativity; Proj. = Projects. * $p <.05$., ** $p <.01$.

ratings and novelty and usefulness ratings), it was notable that no personality measures were significantly correlated with mean ratings for novelty, usefulness, or creativity ($p > .126$ in all cases). Within personality measures, social risk-taking was robustly correlated with intellect ($r = .60$, $p = .001$) and weakly correlated with openness ($r = .19$, $p = .048$). Meanwhile, general risk-taking was weakly but not significantly correlated with intellect ($r = .18$, $p = .063$), and did not correlate with openness ($r = −.04$, $p = .709$).

Single-subject maximum likelihood estimation

SSMLE was conducted for AUT ideas and Projects separately, to estimate standardized coefficients for novelty and usefulness for each participant individually. Differences between coefficients within and across contexts, and relationships between coefficients and personality measures were then examined. Since the statistical significance of individual participant estimates is not of interest to this study, significance values for individual estimates are not included here.

Across all participants, within AUT ideas, coefficients had a mean of 0.48 ($SD = 0.33$) for novelty, and 0.08 ($SD = 0.33$) for usefulness. With Projects, coefficients

had a mean of 0.38 ($SD = 0.39$) for novelty, and 0.24 ($SD = 0.38$) for usefulness. Boxplots summarizing the distributions of these coefficients are presented in Figure 2.

A series of between-participants t-tests were conducted to test for significant differences between novelty and usefulness coefficients, both within and between idea contexts (AUT ideas and Projects). In all t-test results, we report Cohen's d_{av} as a measure of effect size (Lakens, 2013). Results are summarized in Table 2. Novelty coefficients were significantly larger than usefulness coefficients among both AUT ratings and Projects ratings. Comparing across idea context, novelty coefficients were significantly higher among AUT ratings than Projects ratings. By contrast, usefulness coefficients were significantly higher among Projects ratings than AUT ratings. Together, results suggest that novelty plays a greater role in evaluations of creativity than usefulness in both contexts, but is more important in the context of AUT ideas. In addition, results indicate that usefulness is far more important to evaluations of creativity among urban planning projects than among AUT ideas.

Next, we examined whether the weightings given to novelty and usefulness relate to aspects of participants'
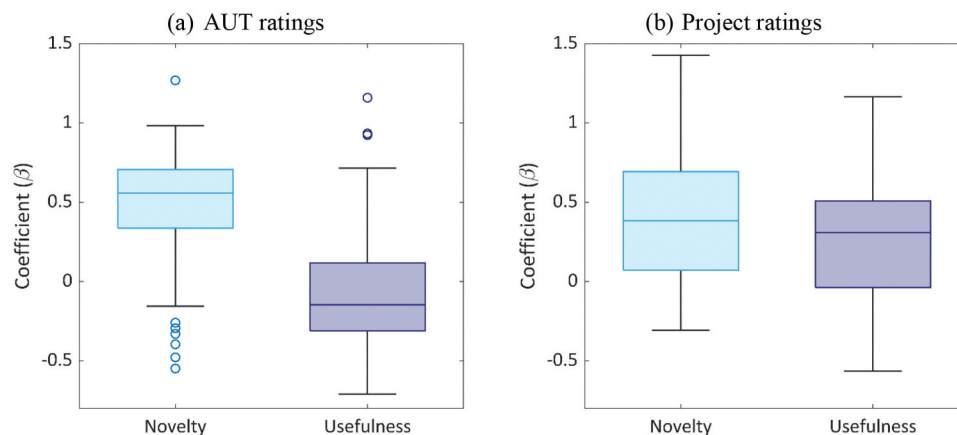

Figure 2. Boxplots showing means and ranges for standardized novelty and usefulness coefficient estimates across all participants, for AUT ratings (a), and Projects ratings (b).

Table 2. Results of t-tests comparing novelty and usefulness coefficients within and between task types.

	t (d.f.)	p	Cohen's d_{av}
AUT: Novelty β > Usefulness β	13.63 (111)	.000	1.29
Projects: Novelty β > Usefulness β	2.35 (110)	.020	0.22
Novelty β: AUT > Projects	2.39 (110)	.018	0.23
Usefulness β: AUT < Projects	7.43 (110)	.000	0.71

Table 3. Correlations between novelty and usefulness coefficient estimates and personality scores.

Task	Coefficient	Openness	Intellect	Risk General	Risk Social
AUT	Novelty β	.23*	.22*	−.03	−.01
	Useful β	.04	.09	.03	−.03
Projects	Novelty β	−.03	−.15	−.14	−.03
	Useful β	.20*	.16	.01	.02

Note. Risk General = general risk-taking; Risk Social = social risk-taking.
*p < .05.

personalities. Correlations between participant personality scores and novelty and usefulness coefficients, for ideas in both contexts, are shown in Table 3.

Among AUT ratings, novelty coefficients were significantly and positively correlated with both openness ($r = .23$, $p = .015$) and intellect ($r = .22$, $p = .023$) scores, while no significant relationships were found between usefulness coefficients and any personality measures ($p > .344$ in all cases). By contrast, among Project ratings, participants' novelty coefficients were not significantly related to any personality measures ($p > .128$ in all cases), while usefulness coefficients were significantly and positively correlated with openness score ($r = .20$, $p = .033$), and positively but non-significantly related to intellect score ($r = .16$, $p = .085$). Together, results suggest that participants' openness and intellect scores may differently moderate the contributions of novelty and usefulness to evaluations of creativity depending on the context (see Figure 3). Specifically, when evaluating the

creativity of AUT ideas, those higher in openness and intellect may place more weight on novelty, while when evaluating the creativity of urban planning projects, the same participants may place more weight on usefulness.

Notably, no measures of risk-taking were found to be significantly related to either novelty or usefulness coefficients, either among AUT ideas or Projects ($p > .159$ in all cases). Therefore, risk-taking measures were left out of subsequent LMEM analyses.

Linear mixed-effects models

The first LMEM (Model 1) was primarily a sanity check to confirm the results of the SSMLE analyses, which had found significant differences between novelty and usefulness coefficients across idea context. As such, this model aimed to test whether idea context significantly moderated the relationships between creativity and novelty and usefulness. In addition, two further LMEMs were constructed to examine AUT ideas (Model 2) and Projects (Model 3) separately. These models were identical in structure and aimed to examine the relative contributions of novelty and usefulness to creativity, the significance of the interaction between novelty and usefulness, and the significance of interactions between openness and intellect and novelty and usefulness. Predictors for novelty, usefulness, openness and intellect (and their interactions) were added successively. Due to multicollinearity concerns, openness and intellect were added separately in the final step of both Model 2 and Model 3.

Model 1 (examining the effect of context) began with a null model containing only random intercepts across participants, with no fixed effects. Following this, main effects for novelty and usefulness were added (Model
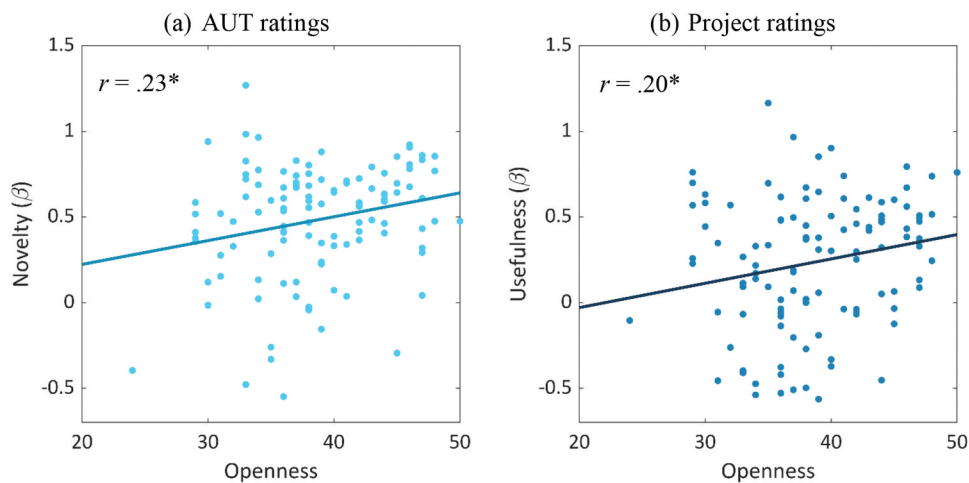

Figure 3. Scatterplots of the relationships between openness and novelty coefficients among AUT ratings (a) and between openness and usefulness coefficients among Project ratings (b). * p < .05.

1A) before random slopes for novelty and usefulness were added (Model 1B). Finally, the main effect for context was added together with interactions between context and novelty and usefulness (Model 1C).

Results are presented in Table 4. Comparing Model 1A to the null model, adding main effects for both novelty and usefulness improved model fit, as indicated by likelihood ratio testing and information criteria. Significant effects were found for both novelty (β = 0.53, SE = 0.01, p < .001), and usefulness (β = −0.10, SE = 0.01, p < .001). Adding random effects slopes for novelty and usefulness in Model 1B also improved model fit, confirming that novelty and usefulness contribute differently to creativity across participants. Finally, in Model 1C, adding effects for context again improved model fit significantly. The main effect of context was not significant (p > .999), while interactions between novelty and context (β = 0.10, SE = 0.03, p < .001), and usefulness and context (β = −0.35, SE = 0.03, p < .001), were highly significant. These results confirm that novelty and usefulness had different relationships with creativity across contexts. Indeed, since AUT ideas were coded as 1 and Projects as 0, the positive moderation effect of Novelty x Context reflects the fact that novelty was more related to creativity among AUT ideas than Projects. Conversely, the negative and larger moderation effect of Usefulness x Context reflects the fact that usefulness was far more related to creativity among Projects than AUT ideas. These results are completely consistent with the SSMLE results (see Table 2 above).

For Model 2 (examining AUT ideas), following the null model (which again contained only random intercepts across participants), main effects for novelty and usefulness were added (Model 2A) before random slopes for novelty and usefulness were added (Model 2B). Next, an interaction effect between novelty and usefulness was added (Model 2C) before main effects and interactions were added for openness (Model 2D.1) and intellect (Model 2D.2) separately.

Results are presented in Table 5. As expected, adding main effects for novelty and usefulness improved the fit of Model 2A relative to the null model. Adding random effects slopes for novelty and usefulness in Model 1B also improved model fit. Significant main effects were found for both novelty (β = 0.50, SE = 0.03, p < .001), and usefulness (β = −0.11, SE = 0.03, p < .001). In Model 2C, a significant interaction was found between novelty and usefulness (β = 0.08, SE = 0.01, p < .001), and this added effect again improved model fit. This result is broadly comparable to previous research that has examined ratings for AUT ideas (Diedrich et al., 2015), which found that usefulness was less related to creativity among non-novel (i.e., common) ideas, and more related to creativity among novel ideas. However, in our data, usefulness was found to be negatively related to creativity among non-novel AUT ideas, while being unrelated to creativity among novel AUT ideas (see Figure 4). Comparing Models 2D.1 and 2D.2 to Model 2C, neither openness nor intellect significantly improved model fit. Main effects for openness and intellect were non-significant (p > .583 in all cases), however significant interactions

Table 4. Linear Mixed-Effects Model (LMEM) of creativity ratings for AUT ideas and projects together, with predictor estimates for novelty, usefulness, and context and interactions.

	Null Model		Model 1A		Model 1B		Model 1C	
Fixed effects	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)
Intercept	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.00	(0.02)
N			0.53	(0.01)***	0.53	(0.03)***	0.40	(0.03)***
U			−0.10	(0.01)***	−0.05	(0.02)	0.23	(0.03)***
Context							0.00	(0.02)
N x Context							0.10	(0.03)***
U x Context							−0.35	(0.03)***
Random Effects	s^2		s^2		s^2		s^2	
Intercept	0.00		0.00		0.00		0.00	
N					0.24		0.22	
U					0.23		0.23	
Model Comparison								
AIC	18185.55		15,624.00		15,624.00		14,598.48	
BIC	18205.88		15,657.89		15,657.89		14,686.58	
$R^2(m)$	.00		.33		.33		.46	
$\Delta\chi^2$ (df)			2565.55 (5)***		2565.55 (10)***		172.61 (13)***	

Note. N = Novelty; U = Usefulness; Context = AUT ideas (coded as 1) vs Projects (coded as 0); Results for fixed effects are presented as standardized regression coefficients with standard error in parentheses; s^2 is the standard error estimate for random intercepts and slopes; Model 1A is compared to the null model, Model 1B is compared to Model 1A, Model 1C is compared to Model 1B; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; $R^2(m)$ = proportion of variation explained by fixed effects (Nakagawa & Schielzeth, 2013); $\Delta\chi^2$ = Likelihood ratio test statistic for comparison of models.
***p < .001.

Table 5. Linear Mixed-Effects Model (LMEM) of creativity ratings for AUT ideas, with predictor estimates for novelty, usefulness, and personality factors and interactions.

	Null Model		Model 2A		Model 2B		Model 2C		Model 2D.1		Model 2D.2	
Fixed effects	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)
Intercept	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.04	(0.01)***	0.04	(0.01)***	0.04	(0.01)***
N			0.49	(0.01)***	0.50	(0.03)***	0.49	(0.03)***	0.49	(0.03)***	0.49	(0.03)***
U			−0.17	(0.01)***	−0.11	(0.03)***	−0.10	(0.03)***	−0.10	(0.03)***	−0.10	(0.03)***
N x U							0.08	(0.01)***	0.08	(0.01)***	0.08	(0.01)***
O									0.00	(0.01)		
N x O									0.07	(0.03)*		
U x O									0.03	(0.03)		
I											0.01	(0.01)
N x I											0.07	(0.03)*
U x I											0.04	(0.03)
Random Effects	s^2		s^2		s^2		s^2		s^2		s^2	
Intercept	0.00		0.00		0.00		0.02		0.02		0.02	
N					0.27		0.27		0.26		0.26	
U					0.28		0.28		0.28		0.27	
Model Comparison												
AIC	15149.24		12,752.45		11,846.56		11,793.32		11,793.48		11,792.93	
BIC	15169.01		12,785.40		11,912.45		11,865.81		11,885.74		11,885.18	
$R^2(m)$	.00		.36		.50		.50		.50		.50	
Δχ² (df)			2400.79 (5)***		915.90 (10)***		55.24 (11)***		5.84 (14)		6.39 (14)	

Note. N = Novelty; U = Usefulness; O = Openness; I = Intellect; Results for fixed effects are presented as standardized regression coefficients with standard error in parentheses; s^2 is the standard error estimate for random intercepts and slopes; Model 2A is compared to the null model, Model 2B is compared to Model 2A, Model 2C is compared to Model 2B, and Model 2D.1 and 2D.2 are each compared to Model 2C; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; $R^2(m)$ = proportion of variation explained by fixed effects (Nakagawa & Schielzeth, 2013). Δχ² = Likelihood ratio test statistic for comparison of models.
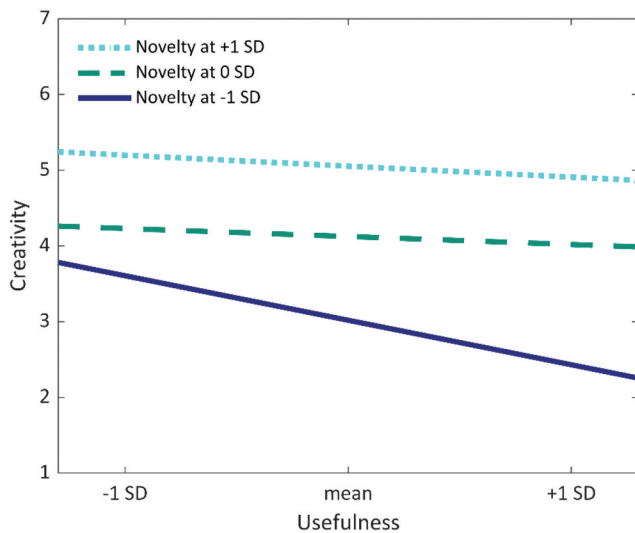*p < .05; *** p < .001.

Figure 4. Simple slopes plot of the interaction between novelty and usefulness as predictors of creativity among AUT ratings.

were found between both novelty and openness (β = 0.07, SE = 0.03, p = .015; see Figure 5a) and novelty and intellect (β = 0.07, SE = 0.03, p = .014), suggesting that both aspects of Openness/Intellect lead to a greater consideration of novelty when participants evaluated the creativity of AUT ideas. Neither openness nor intellect interacted significantly with usefulness (p > .185 in all cases).

Model 3 was constructed in exactly the same way as Model 2 but now applied in the context of Projects

rather than AUT ideas. Results are presented in Table 6. Relative to the null model, adding main effects for novelty and usefulness again improved the fit of Model 3A. Adding random effects slopes for novelty and usefulness in Model 3B also improved model fit. Significant main effects were found for both novelty (β = 0.41, SE = 0.03, p < .001), and usefulness (β = 0.22, SE = 0.03, p < .001). In Model 3C, an interaction between novelty and usefulness was added, but this was non-significant (β = −0.04, SE = 0.03, p = .116), and did not improve model fit. Comparing Models 3D.1 and 3D.2 to Model 3C, neither openness nor intellect significantly improved model fit. Main effects for openness and intellect were non-significant (p > .964 in all cases). A significant interaction was found between usefulness and openness (β = 0.08, SE = 0.03, p = .020), while an interaction between usefulness and intellect did not reach significance (β = 0.06, SE = 0.03, p = .063). These results suggest that participants who are higher in openness may place greater importance on usefulness when evaluating the creativity of Projects (see Figure 5b). Neither openness nor intellect interacted significantly with novelty (p > .155 in all cases).

Discussion

If creative ideas are both novel and useful (Runco & Jaeger, 2012; Stein, 1953), individuals should weigh up
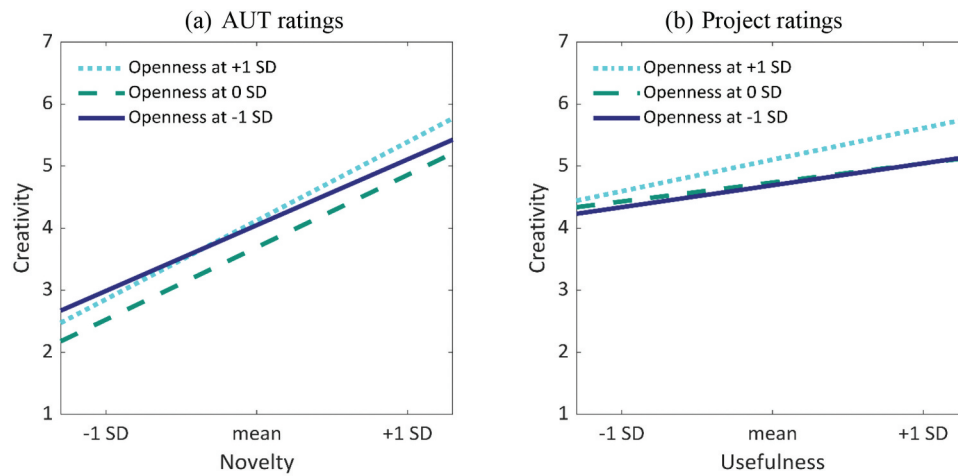
Figure 5. Simple slopes plot of the interaction between openness and novelty, among AUT ratings (a), and between openness and usefulness, among Project ratings (b).

Table 6. Linear Mixed-Effects Model (LMEM) of creativity ratings for projects, with predictor estimates for novelty, usefulness, and personality factors and interactions.

	Null Model		Model 3A		Model 3B		Model 3C		Model 3D.1		Model 3D.2	
Fixed effects	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)	β	(SE_b)
Intercept	0.00	(0.03)	0.00	(0.02)	0.00	(0.02)	0.01	(0.02)	0.01	(0.02)	0.01	(0.02)
N			0.41	(0.03)***	0.40	(0.03)***	0.39	(0.03)***	0.39	(0.03)***	0.39	(0.03)***
U			0.22	(0.03)***	0.23	(0.03)***	0.23	(0.03)***	0.23	(0.03)***	0.23	(0.03)***
N x U							−0.04	(0.03)	−0.04	(0.03)	−0.04	(0.03)
O									0.00	(0.02)		
N x O									−0.01	(0.03)		
U x O									0.08	(0.03)*		
I											0.00	(0.02)
N x I											−0.05	(0.03)
U x I											0.06	(0.03)
Random Effects	s^2		s^2		s^2		s^2		s^2		s^2	
Intercept	0.00		0.00		0.00		0.00		0.00		0.00	
N					0.22		0.22		0.22		0.21	
U					0.24		0.24		0.23		0.23	
Model Comparison												
AIC	3056.37		2712.51		2680.53		2680.06		2680.25		2682.07	
BIC	3071.44		2737.62		2730.74		2735.29		2750.54		2752.37	
$R^2(m)$	.00		.27		.35		.35		.35		.35	
$\Delta\chi^2$ (df)			347.86 (5)***		41.99 (10)***		2.47 (11)		5.81 (14)		3.99 (14)	

Note. N = Novelty; U = Usefulness; O = Openness; I = Intellect; Results for fixed effects are presented as standardized regression coefficients with standard error in parentheses; s^2 is the standard error estimate for random intercepts and slopes; Model 3A is compared to the null model, Model 3B is compared to Model 3A, Model 3C is compared to Model 3B, and Model 3D.1 and 3D.2 are both compared to Model 3C; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; $R^2(m)$ = proportion of variation explained by fixed effects (Nakagawa & Schielzeth, 2013); $\Delta\chi^2$ = Likelihood ratio test statistic for comparison of models.
*$p < .05$; *** $p < .001$.

these two components when evaluating the creativity of ideas. The present study focused on the evaluation of exogenous (i.e., non-self-generated) ideas, and examined how the weightings applied to novelty and usefulness vary according to the context of the idea and the personality of the rater. Both SSMLE and LMEM analyses indicated that the relative importance of novelty and usefulness to evaluations of creativity can vary widely over different contexts, and that those with different personalities may consider novelty and usefulness to different extents. Specifically, while novelty was more

important to evaluations of creativity than usefulness among both AUT ideas and Projects, we found that usefulness was far more important in the context of Projects than in the context of AUT ideas. Moreover, we found that individuals higher in openness (and to a lesser extent, intellect) placed a greater emphasis on novelty when evaluating AUT ideas, while placing a greater emphasis on usefulness when evaluating Projects.

The finding that raters generally consider novelty more than usefulness when evaluating creativity was in

line with our predictions and with prior research (Acar et al., 2017; Caroff & Besançon, 2008; Diedrich et al., 2015; Han et al., 2021; Runco & Charles, 1993). Among both AUT ideas and Projects, novelty coefficients were significantly greater than usefulness coefficients in SSMLE analyses, while LMEMs found larger coefficients for novelty than usefulness in both contexts.

However, we also found clear differences between contexts. In line with predictions, raters considered novelty more in the context of AUT ideas, and usefulness more in the context of Projects. Specifically, LMEMs revealed significant interactions between idea context and both novelty and usefulness, while SSMLE analyses revealed greater novelty coefficients among AUT ideas than Projects, and greater usefulness coefficients among Projects than AUT ideas. Indeed, separate LMEMs for AUT ideas and Projects suggested that while usefulness was negatively related to creativity among AUT ideas, it was positively related to creativity among Projects. These findings are consistent with the notion that different contexts can lead to different considerations of novelty and usefulness (Long, 2014; Runco et al., 2005). In contexts such as the AUT, where ideas are unlikely to be used in the real world, usefulness may not contribute to evaluations of creativity, or may even contribute negatively. By contrast, in contexts where ideas are clearly applicable to the real-world, usefulness may play a far greater role in evaluations of creativity. These results extend previous research by highlighting how the context in which an idea was generated can impact evaluations of creativity.

Considering the role of rater personality in evaluations of creativity, findings were more nuanced than expected. While we predicted that higher openness would be related to a greater consideration of novelty, this was only the case in the context of AUT ideas. In the context of Projects, higher openness was related to a greater consideration of usefulness. In addition, while we suggested that intellect might relate to a greater consideration of usefulness in both contexts, it followed the same context-dependent pattern as openness. Specifically, among AUT ideas, both openness and intellect were positively correlated with SSMLE novelty coefficients, while significant interactions between these traits and novelty were found in an LMEM. By contrast, among Projects, openness and intellect were positively correlated with SSMLE usefulness coefficients (though for intellect this correlation was non-significant), while significant interactions between openness and usefulness were found in an LMEM. Indeed, there was no evidence to suggest that openness and intellect were differently related to considerations of novelty and

usefulness, as might be expected based on the different relationships between these traits and achievements in the arts and sciences (Kaufman et al., 2016). Overall, this suggests that the twin aspects of openness and intellect may be better considered as a single trait in the context of creativity evaluations.

Also contrary to our predictions, we found no significant relationships between risk-taking and novelty and usefulness coefficients, suggesting that an individual's preference for risk-taking does not relate to different considerations of these components when evaluating creativity. However, it is important to consider that the present study examined the evaluation of exogenous ideas. If participants had instead evaluated the creativity of their own ideas, it is plausible that their risk-taking preference, and indeed their openness/intellect score, might have a more profound impact on their consideration of novelty and usefulness (see Rodriguez et al., 2020; Silvia, 2008).

Considering other findings, in line with previous research (Diedrich et al., 2015) we found an interaction between novelty and usefulness among AUT ideas. Examination of a simple slopes plot of this interaction (see Figure 4) indicates that, in our study, usefulness was negatively related to creativity among non-novel ideas and unrelated to creativity among novel ideas. By contrast, we found no such interaction among Projects. It was also notable that openness and intellect did not show significant main effects as predictors of creativity in the LMEMs, either among AUT ideas or among Projects. Indeed, no significant correlations were found between any personality measures and participant mean ratings for creativity, novelty, or usefulness in either context. This suggests that while the trait openness/intellect plays a role in how raters weigh novelty and usefulness, it may not impact overall creativity judgments, at least in the case of exogenous ideas.

Impact and implications

To our knowledge, the present study is the first to investigate how differences in idea context and rater personality can lead to different considerations of novelty and usefulness during evaluations of creativity. Our findings help extend a growing body of work that has examined how individuals consider novelty and usefulness when evaluating creativity (Acar et al., 2017; Caroff & Besançon, 2008; Diedrich et al., 2015; Runco & Charles, 1993; Storme & Lubart, 2012), and how variations in the evaluation of creativity relate to individual differences (Herman & Reiter-Palmon, 2011; Karwowski et al., 2020; Lee et al., 2017; Mastria et al., 2019; Mueller

et al., 2012). Overall, our findings highlight the importance of considering contextual and interpersonal factors when researchers examine how creativity is evaluated, defined, and perceived, strengthening recent calls for creativity assessments that can account for variation across raters (Barbot et al., 2019; Myszkowski & Storme, 2019). Indeed, it seems likely that both the generation and evaluation of creative ideas may involve markedly different processes depending on both the individual in question and the context of the problem. Different individuals may consider different criteria more important than others when performing creative tasks and may use a different balance of cognitive processes to produce ideas that meet these criteria. Similarly, different creative contexts may call for different levels of novelty and usefulness (or other components), leading individuals to weigh these aspects differently when they evaluate ideas depending on the specific requirements of the problem.

Limitations and future directions

One limitation of the present research is that it did not assess the intelligence or creativity of the raters. Intelligence has been linked to a greater consideration of novelty when raters evaluate creativity (Storme & Lubart, 2012), and it would be interesting to examine how intelligence interacts with the consideration of novelty and usefulness in different contexts. For example, intelligence might follow a similar pattern to openness, relating to a greater consideration of novelty among AUT ideas and greater consideration of usefulness among real-world projects. Meanwhile, assessing creativity would allow researchers to better examine links between how individuals generate their ideas and how they evaluate the ideas of others. For example, do individuals who tend to generate highly novel but nonuseful ideas themselves also consider novelty more than usefulness when evaluating the ideas of others? These questions should be examined by future research.

Indeed, assessing the creativity of raters (e.g., by having them complete the AUT) would also provide an opportunity for them to evaluate their own ideas. The present study focused on the evaluation of exogenous ideas which, while more relevant to creativity assessment methodologies, has been found to differ from the evaluation of self-generated ideas (Karwowski et al., 2020; Rodriguez et al., 2020; Runco & Smith, 1992). It is possible that individuals consider novelty and usefulness differently when evaluating the creativity of their own ideas as opposed to others' ideas. It is also possible that personality traits play a different role depending on whether participants evaluate their own ideas or others'

ideas. For example, research has found that individuals with higher general personality scores provide higher-quality evaluations of exogenous ideas, but lower-quality evaluations of their own ideas (Rodriguez et al., 2020). As such, future studies could examine and compare evaluations of both self-generated and exogenous ideas.

Moreover, the present study focused on only the two most widely discussed components of creativity: novelty and usefulness. However, research suggests that additional factors, such as surprise (Acar et al., 2017; Simonton, 2018), may also be considered by individuals when they evaluate creativity. Indeed, the best-fitting LMEMs in the present study only explained around 50% of the variance in creativity ratings, indicating considerable room for other explanatory factors. Future studies could therefore collect additional ratings for other components of creativity.

A further option for future studies is to examine relationships between mood and uncertainty and the weightings placed on novelty and usefulness. Indeed, prior research has indicated that more positive moods (Mastria et al., 2019), and more certainty among raters (Lee et al., 2017), relate to higher creativity ratings of exogenous ideas, while greater promotion focus is related to more accuracy when evaluating the novelty of one's own ideas (Herman & Reiter-Palmon, 2011). Together, this research implies that some individuals may show a greater affinity for creative and novel ideas, and led us to expect that those with higher openness and risk-taking scores might place a greater emphasis on novelty when evaluating creativity. However, the relationships between openness and considerations of novelty and usefulness were found to depend on the context, while no relationships were found for risk-taking. Future research could assess or manipulate the promotion vs. prevention focus of raters, as well as their current mood and level of certainty, to examine how these factors specifically influence considerations of novelty and usefulness. For example, does greater uncertainty lead to a greater consideration of usefulness when participants evaluate creative ideas?

Conclusion

Examining differences in how individuals evaluate creativity, and the factors they consider during their evaluations, is an emerging area of research with relatively few existing studies. In the present study, we found that both the context of ideas and the personality of raters play important and interacting roles in how novelty and usefulness are considered in evaluations of creativity. There is enormous potential for further research to investigate the factors (including mood, personality, intelligence, and cultural background) that can influence how individuals weigh up different aspects of an idea when assessing its creativity.

After all, evaluation is a critical part of creative cognition. Understanding how creativity is perceived and defined in different contexts and across different raters is highly important not just to our understanding of subjective assessments, but to our understanding of creativity itself.

Acknowledgments

Disclosure statement

ORCID

James Lloyd-Cox 🔟 http://orcid.org/0000-0003-3492-9638
Alan Pickering 🔟 http://orcid.org/0000-0002-7301-5321
Joydeep Bhattacharya 🔟 http://orcid.org/0000-0003-3443-9049

References

Acar, S., Berthiaume, K., Grajzel, K., Dumas, D., & Flemister, C., "Tedd," & Organisciak, P. (2021 December) Applying automated originality scoring to the verbal form of Torrance tests of creative thinking. *Gifted Child Quarterly*, 001698622110618. 10.1177/00169862211061874

Acar, S., Burnett, C., & Cabra, J. F. (2017). Ingredients of creativity: Originality and more. *Creativity Research Journal*, *29*(2), 133–144. doi:10.1080/10400419.2017. 1302776

Acar, S., Runco, M. A., & Park, H. (2019). What should people be told when they take a divergent thinking test? A meta-analytic review of explicit instructions for divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 39–49. doi:10.1037/aca0000256

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*(5), 997–1013. doi:10.1037/0022-3514.43.5.997

Baer, J., & McKool, S. S. (2014). The gold standard for assessing creativity. *International Journal of Quality Assurance in Engineering and Technology Education*, *3*(1), 81–93. doi:10.4018/ijqaete.2014010104

Barbot, B., Hass, R. W., & Reiter-Palmon, R. (2019). Creativity assessment in psychological research: (Re)setting the standards. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 233–240. doi:10.1037/ACA0000233

Basadur, M. (1995). Optimal ideation-evaluation ratios. *Creativity Research Journal*, *8*(1), 63–75. doi:10.1207/s15326934crj0801

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, *132*(4), 355–429. doi:10.3200/MONO.132.4.355-430

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780. doi:10.3758/s13428-020-01453-w

Blais, A., & Weber, E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgement and Decision Making*, *1*(1), 33–47.

Boden, M. A. (2007). Creativity in a nutshell. *Think*, *5*(15), 83–96. doi:10.1017/s147717560000230x

Bonetto, E., Pichot, N., Pavani, J. B., & Adam-Troïan, J. (2021). Creative individuals are social risk-takers: Relationships between creativity, social risk-taking and fear of negative evaluations. *Creativity*, *7*(2), 309–320. doi:10.2478/ctra-2020-0016

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, *67*(6), 380–400. doi:10.1037/H0040373

Caroff, X., & Besançon, M. (2008). Variability of creativity judgments. *Learning and Individual Differences*, *18*(4), 367–371. doi:10.1016/j.lindif.2008.04.001

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, *17*(1), 37–50. doi:10.1207/S15326934CRJ1701_4

Ceh, S. M., Edelmann, C., Hofer, G., & Benedek, M. (2022). Assessing raters: What factors predict discernment in novice creativity raters? *Journal of Creative Behavior*, *56*(1), 41–54. doi:10.1002/jocb.515

Corazza, G. E. (2016). Potential originality and effectiveness: The dynamic definition of creativity. *Creativity Research Journal*, *28*(3), 258–267. doi:10.1080/10400419.2016.1195627

Cropley, D. H., & Marrone, R. L. (2021). Automated scoring of figural creativity using a convolutional neural network. *PsyArXiv, (Online pre-print)*. doi:10.31234/OSF.IO/8QE7Y

Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 159–166. doi:10.1037/ACA0000220

Dewett, T. (2007). Linking intrinsic motivation, risk taking, and employee creativity in an R&D environment. *R and D Management*, *37*(3), 197–208. doi:10.1111/j.1467-9310.2007.00469.x

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*(4), 825–858. doi:10.1111/J.1467-6494.2005.00330.X

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the big five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. doi:10.1037/0022-3514.93.5.880

Diedrich, J., Benedek, M., Jauk, E., & Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, and the Arts*, *9*(1), 35–40. doi:10.1037/a0038688

Diedrich, J., Jauk, E., Silvia, P. J., Gredlein, J. M., Neubauer, A. C., & Benedek, M. (2018). Assessment of real-life creativity: The inventory of creative activities and achievements (ICAA). *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 304–316. doi:10.1037/ACA0000137

Edwards, K. M., Peng, A., Miller, S. R., & Ahmed, F. (2021). If a picture is worth 1000 words, is a word worth 1000 features for design metric estimation? *Proceedings of the ASME Design Engineering Technical Conference*, 6. doi:10.1115/DETC2021-70158

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, 236(5), 119–127. Retrieved from http://www.jstor.org/stable/24954030

Erbas, A. K., & Bas, S. (2015). The contribution of personality traits, motivation, academic risk-taking and metacognition to the creative ability in mathematics. *Creativity Research Journal*, 27(4), 299–307. doi:10.1080/10400419.2015.1087235

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. doi:10.3758/BF03193146

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4), 290–309. doi:10.1207/s15327957pspr0204_5

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Gibson, C., & Mumford, M. D. (2013). Evaluation, criticism, and creativity: Criticism content and effects on creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts*, 7(4), 314–331. doi:10.1037/A0032616

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology*, 98(4), 611–625. doi:10.1111/J.2044-8295.2007.TB00467.X

Glover, J. A., & Sautter, F. (1977). Relation of four components of creativity to risk-taking preferences. *Psychological Reports*, 41(1), 227–230. doi:10.2466/pr0.1977.41.1.227

Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York, NY: Mc-Graw Hill.

Han, J., Forbes, H., & Schaefer, D. (2021). An exploration of how creativity, functionality, and aesthetics are related in design. *Research in Engineering Design*, 32(3), 289–307. doi:10.1007/s00163-021-00366-9

Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, 61, 569–598. doi:10.1146/annurev.psych.093008.100416

Herman, A., & Reiter-Palmon, R. (2011). The effect of regulatory focus on idea generation and idea evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, 5(1), 13–20. doi:10.1037/a0018587

Ivancovsky, T., Shamay-Tsoory, S., Lee, J., Morio, H., & Kurman, J. (2019). A dual process model of generation and evaluation: A theoretical framework to examine cross-cultural differences in the creative process. *Personality and Individual Differences*, 139, 60–68. July 2018. doi:10.1016/j.paid.2018.11.012

Karwowski, M., Czerwonka, M., & Kaufman, J. C. (2020). Does intelligence strengthen creative metacognition? *Psychology of Aesthetics, Creativity, and the Arts*, 14(3), 353–360. doi:10.1037/ACA0000208

Katahira, K. (2016). How hierarchical models improve point estimates of model parameters at the individual level. *Journal of Mathematical Psychology*, 73, 37–58. doi:10.1016/j.jmp.2016.03.007

Kaufman, J. C. (2019). Self-assessments of creativity: Not ideal, but better than you think. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 187–192. doi:10.1037/ACA0000217

Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, 2(2), 96–106. doi:10.1016/J.TSC.2007.04.002

Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., . . . DeYoung, C. G. (2016). Openness to experience and intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality*, 84(2), 248–258. doi:10.1111/jopy.12156

Kharkhurin, A. V. (2014). Creativity.4in1: Four-criterion construct of creativity. *Creativity Research Journal*, 26(3), 338–352. doi:10.1080/10400419.2014.929424

Kleinmintz, O. M., Goldstein, P., Mayseless, N., Abecasis, D., Shamay-Tsoory, S. G., & Antonietti, A. (2014). Expertise in musical improvisation and creativity: The mediation of idea evaluation. *PLoS ONE*, 9(7), e101568. doi:10.1371/journal.pone.0101568

Kleinmintz, O. M., Ivancovsky, T., & Shamay-Tsoory, S. G. (2019). The twofold model of creativity: The neural underpinnings of the generation and evaluation of creative ideas. *Current Opinion in Behavioral Sciences*, 27, 131–138. doi:10.1016/j.cobeha.2018.11.004

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. doi:10.3389/fpsyg.2013.00863

Lee, Y. S., Chang, J. Y., & Choi, J. N. (2017). Why reject creative ideas? Fear as a driver of implicit bias against creativity. *Creativity Research Journal*, 29(3), 225–235. doi:10.1080/10400419.2017.1360061

Long, H. (2014). More than appropriateness and novelty: Judges' criteria of assessing creative products in science tasks. *Thinking Skills and Creativity*, 13, 183–194. doi:10.1016/j.tsc.2014.05.002

Lubart, T. I. (2001). Models of the creative process: Past, present and future. *Creativity Research Journal*, 13(3–4), 295–308. doi:10.1207/S15326934CRJ1334

Luft, C. D. B., Zioga, I., Thompson, N. M., Banissy, M. J., & Bhattacharya, J. (2018). Right temporal alpha oscillations as a neural mechanism for inhibiting obvious associations. *Proceedings of the National Academy of Sciences of the United States of America*, 115(52), E12144–E12152. doi:10.1073/pnas.1811465115

Martin, L., & Wilson, N. (2017). Defining creativity with discovery. *Creativity Research Journal*, 29(4), 417–425. doi:10.1080/10400419.2017.1376543

Mastria, S., Agnoli, S., Corazza, G. E., & Eisenbarth, H. (2019). How does emotion influence the creativity evaluation of exogenous alternative ideas? *PLOS ONE*, 14(7), e0219298. doi:10.1371/JOURNAL.PONE.0219298

McCarthy, M., Chen, C. C., & McNamee, R. C. (2018). Novelty and usefulness trade-off: Cultural cognitive differences and creative idea evaluation. *Journal of Cross-Cultural Psychology*, 49(2), 171–198. doi:10.1177/0022022116680479

McCrae, R. R., & Ingraham, L. J. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality*

and Social Psychology, *52*(6), 1258–1265. doi:10.1037/0022-3514.52.6.1258

McIntosh, T., Mulhearn, T. J., & Mumford, M. D. (2021). Taking the good with the bad: The impact of forecasting timing and valence on idea evaluation and creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(1), 111–124. doi:10.1037/ACA0000237

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*(1), 20–35. doi:10.1037/MET0000182

Mueller, J. S., Melwani, S., & Goncalo, J. A. (2012). The bias against creativity: Why people desire but reject creative ideas. *Psychological Science*, *23*(1), 13–17. doi:10.1177/0956797611421018

Myszkowski, N., & Storme, M. (2019). Judge response theory? A call to upgrade our psychometrical account of creativity judgments. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 167–175. doi:10.1037/aca0000225

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixedeffects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. doi:10.1111/J.2041-210X.2012.00261.X

Oleynick, V. C., DeYoung, C. G., Hyde, E., Barry Kaufman, S., Beaty, R. E., & Silvia, P. J. (2017). Openness/Intellect: The core of the creative personality. In G. J. Feist, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *Cambridge handbook of creativity and personality research* (pp. 9–27). New York: Cambridge University Press.

Pétervári, J. (2018). *The evaluation of creative ideas – Analysing the differences between expert and novice judges.* (Doctoral Thesis): Queen Mary University of London, London. Retrieved from http://qmro.qmul.ac.uk/xmlui/handle/123456789/42526.

Pinho, A. L., de Manzano, Ö., Fransson, P., Eriksson, H., & Ullén, F. (2014). Connecting to create: Expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *Journal of Neuroscience*, *34*(18), 6156–6163. doi:10.1523/JNEUROSCI.4769-13.2014

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*(2), 83–96. doi:10.1207/s15326985ep3902_1

Plucker, J., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *Journal of Creative Behavior*, *45*(1), 1–22. doi:10.1002/j.2162-6057.2011.tb01081.x

Prabhakaran, R., Green, A. E., & Gray, J. R. (2014). Thin slices of creativity: Using single-word utterances to assess creative cognition. *Behavior Research Methods*, *46*(3), 641. doi:10.3758/S13428-013-0401-7

Puryear, J. S., & Lamb, K. N. (2020). Defining creativity: How far have we come since Plucker, Beghetto, and Dow? *Creativity Research Journal*, *32*(3), 206–214. doi:10.1080/10400419.2020.1821552

Rodriguez, W. A., Cheban, Y., Shah, S., & Watts, L. L. (2020). The general factor of personality and creativity: Diverging effects on intrapersonal and interpersonal idea evaluation.

Personality and Individual Differences, *167*, 110229. doi:10.1016/j.paid.2020.110229

Rominger, C., Papousek, I., Perchtold, C. M., Weber, B., Weiss, E. M., & Fink, A. (2018). The creative brain in the figural domain: Distinct patterns of EEG alpha power during idea generation and idea elaboration. *Neuropsychologia*, *118*, 13–19. doi:10.1016/J.NEUROPSYCHOLOGIA.2018.02.013

Runco, M. A., & Charles, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences*, *15*(5), 537–546. doi:10.1016/0191-8869(93)90337-3

Runco, M. R., Illies, J. J., & Eisenman, R. (2005). Creativity, originality, and appropriateness: What do explicit instructions tell us about their relationships? *Journal of Creative Behavior*, *39*(2), 137–148. doi:10.1002/j.2162-6057.2005.tb01255.x

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92–96. doi:10.1080/10400419.2012.650092

Runco, M. A., & Smith, W. R. (1992). Interpersonal and intrapersonal evaluations of creative ideas. *Personality and Individual Differences*, *13*(3), 295–302. doi:10.1016/0191-8869(92)90105-X

Shen, W., Hommel, B., Yuan, Y., Chang, L., & Zhang, W. (2018). Risk-taking and creativity: Convergent, but not divergent thinking is better in low-risk takers. *Creativity Research Journal*, *30*(2), 224–231. doi:10.1080/10400419.2018.1446852

Silvia, P. J. (2008). Discernment and creativity: How well can people identify their most creative ideas? *Psychology of Aesthetics, Creativity, and the Arts*, *2*(3), 139–146. doi:10.1037/1931-3896.2.3.139

Simonton. (1999). *Origins of genius. Darwinian perspectives on creativity.* Oxford: NY: Oxford University Press.

Simonton, D. K. (2013). Creative thought as blind variation and selective retention: Why creativity is inversely related to sightedness. *Journal of Theoretical and Philosophical Psychology*, *33*(4), 253–266. doi:10.1037/a0030705

Simonton, D. K. (2018). Defining creativity: Don't we also need to define what is not creative? *Journal of Creative Behavior*, *52*(1), 80–90. doi:10.1002/jocb.137

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In *New methods in cognitive psychology* (pp. 4–31). Taylor and Francis. doi:10.4324/9780429318405-2

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. doi:10.1177/1745691616658637

Stein, M. I. (1953). Creativity and culture. *The Journal of Psychology*, *36*(2), 311–322. doi:10.1080/00223980.1953.9712897

Sternberg, R. J. (2018). A triangular theory of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(1), 50–67. doi:10.1037/aca0000095

Storme, M., & Lubart, T. (2012). Conceptions of creativity and relations with judges' intelligence and personality. *Journal of Creative Behavior*, *46*(2), 138–149. doi:10.1002/jocb.10

Taylor, C. W. (1988). Various approaches to and definitions of creativity. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 99–121). Cambridge University Press.

Treffinger, D. (1992). Stimulating creativity: Issues and future directions. In S. G. Isaksen, M. C. Murdock, R. L. Firestein, & D. J. Treffinger (Eds.), *Nurturing and developing creativity: The emergence of a discipline* (pp. 8–27). Norwood, NJ: Ablex.

Tyagi, V., Hanoch, Y., Hall, S. D., Runco, M., & Denham, S. L. (2017). The risky side of creativity: Domain specific risk taking in creative individuals. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.00145

Watts, L. L., Steele, L. M., Medeiros, K. E., & Mumford, M. D. (2019). Minding the gap between generation and implementation: Effects of idea source, goals, and climate on selecting and refining creative ideas. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(1), 2–14. doi:10.1037/ACA0000157

Wilken, A., Forthmann, B., & Holling, H. (2020). Instructions moderate the relationship between creative performance in figural divergent thinking and reasoning capacity. *Journal of Creative Behavior*, *54*(3), 582–597. doi:10.1002/jocb.392

Wilson, R. C., Guilford, J. P., & Christensen, P. R. (1953). The measurement of individual differences in originality. *Psychological Bulletin*, *50*(5), 362–370. doi:10.1037/H0060857