Psychology of Aesthetics, Creativity, and the Arts

Incongruent Objects Are Surprising But Do Not Inspire More Original Ideas

Wendy Ross, Sergio Agnoli, and Vlad Glăveanu

Online First Publication, December 11, 2025. https://dx.doi.org/10.1037/aca0000829

AMERICAN PSYCHOLOGICAL ASSOCIATION
DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

ISSN: 1931-3896
https://doi.org/10.1037/aca0000829

Incongruent Objects Are Surprising But Do Not Inspire More Original Ideas

Wendy Ross[1], Sergio Agnoli[2], and Vlad Glăveanu[3]
[1] Department of Psychology, London Metropolitan University
[2] Department of Life Sciences, University of Trieste
[3] Department of Psychology, Dublin City University

This study investigates the relationship between surprise and creative ideation through four preregistered experiments (total $N = 964$). In Experiments 1a ($N = 117$) and 1b ($N = 143$), participants viewed images of objects with differing levels of congruence with the environmental context using stimuli from the Scene Grammar (SCEGRAM) database, then completed an Alternative Uses Task for those objects. Experiment 2 ($N = 316$) replicated these findings using the ObjAct stimulus set, which depicted humans manipulating congruent versus incongruent objects. Experiment 3 ($N = 388$) controlled for order effects and stimulus believability using a reduced stimulus set. While the experimental manipulations successfully elicited different levels of reported surprise, this did not translate to increased originality in uses for the objects. Across all experiments, there was no significant relationship between ratings of surprise and the originality of ideas produced. The findings suggest that simple schema disruption or expectancy violation is not sufficient to enhance creative thinking, at least in the context of an online experimental task. Potential explanations for the null effect have implications for wider theories of surprise, including the distinction between assessing something as surprising (cognitive surprise) versus feeling surprise (affective surprise), the role of agency and personal significance in surprise experiences, and the possibility that surprise may play a smaller role in creative tasks than theorized.

Keywords: surprise, creativity, originality, inspiration, Alternative Uses Task

How a new thought is generated is a key question for research in higher level cognitive processes (Green et al., 2023). When presented with a challenge that requires a novel solution, the human agent's first reaction will often be to draw on prior experience (Gilhooly et al., 2007; Mednick, 1962). While this is an important part of creative ideation, there is some evidence that prior experience may block the generation of truly novel ideas. Famously, Ward showed that when asked to generate ideas for aliens, people stayed remarkably close to a humanoid form, with bilateral symmetry, sensory receptors, and limb-like appendages (Ward, 1994). Creativity requires moving beyond what is already known to address new possibilities that are not simply reproductions of prior experience (Corazza, 2023; Ohlsson, 2011). To do this, people can generate new ideas through internal processes (Benedek et al., 2018) or they can encounter new situations that seed novel possibilities (Hanchett Hanson, 2023). These new situations are often described as being surprising; however, the empirical evidence for a direct link between surprise and creative ideation remains limited and inconsistent. We present an empirical investigation into whether a surprising stimulus—specifically, an object placed in an incongruent visual context—can increase the originality of ideas people generate about it.

Moments of external triggers for inspiration seem particularly important to scientific discovery, where experiments can reveal something unexpected and have been theorized as the explanation for breakthrough thinking (Perkins, 2001). Scientists are exhorted to find ways to allow themselves to be open to being surprised, and surprise is even lauded as an epistemic virtue (Currie, 2018; French & Murphy, 2023). Philosophical discussion on the value of surprise to discovery filters through to folk ideas of the importance of being open to being surprised by the unexpected as a way of coming up with a new idea. Such views are often referenced in online advice (Brito, 2021; Brunelle, 2021), accompanied by activities designed to induce surprise or break routines as ways to overcome these cognitive constraints and foster more flexible, original thinking (see, e.g., the guide to leadership, Williams, 2015). It is therefore important that the ongoing folk assumptions about the relationship between surprise and creativity are investigated.

Selcuk Acar served as action editor.
Wendy Ross https://orcid.org/0000-0002-0461-7660
The data for Experiments 1a and 1b were first presented at the Experimental Psychology Society meeting in Plymouth, April 2023, and we would like to thank the audience for helpful feedback. Thomas C. Ormerod also gave helpful comments on the analysis and a first draft of the article. This research was funded by Webster University, Geneva, through a grant awarded to Vlad Glăveanu and Wendy Ross, where Vlad Glăveanu was previously affiliated. The authors have no conflicts of interest to declare. The registrations, data, and the analysis code for all reported experiments can be found at https://osf.io/4n8tc/?view_only=5f82eebeeac347b6a084f28a0356109a. The study was granted institutional approval on October 24, 2022, by London Metropolitan University, Reference SSSP-2022-0001. All registrations, code, and data can be found on the Open Science Framework at https://osf.io/4n8tc/?view_only=5f82eebeeac347b6a084f28a0356109a.

Correspondence concerning this article should be addressed to Wendy Ross, Department of Psychology, London Metropolitan University, Tower Building, Holloway Road, London N7 8DB, United Kingdom. Email: w.ross@londonmet.ac.uk

The Relationship Between Surprise and Ideation

Surprise plays a crucial role in interrupting ongoing mental processes, effectively halting current trains of thought and redirecting attentional resources toward the unexpected event. There are two potential cognitive mechanisms underlying the relationship between having an original thought and experiencing surprise. The traditional view is that the feeling of surprise is valuable because it indicates the disruption of existing mental schemas and the reorientation of attention (Meyer et al., 1991). As Boden (2004, p. 41) wrote, "to be surprised is to find that some of one's previous expectations do not fit the case." From this perspective, the effect of surprise is closely related to breaking free from prior conceptions and constraints. This understanding of surprise aligns with research that suggests that a broadened attentional ability toward seemingly irrelevant information, if processed during idea generation, leads to more original ideas (Agnoli et al., 2015; Mastria et al., 2022). It also aligns with the understanding that the feeling of insight (of which surprise is hypothesized to be a major component; Danek et al., 2014) reflects a rapid cognitive restructuring (Danek et al., 2020). In other words, the value of a surprising event is that it introduces novelty into the cognitive system (Ross, 2024; Ross & Arfini, 2023).

While these views of surprise as a reaction to schema discrepancy or rapid restructuring address the initial perceptual unexpectedness that occurs prior to experiencing surprise, recent theories have placed a greater emphasis on the "sense-making" aspects of a surprising event, that is, the resolution of the doubt that a surprising event has caused. This recalls the role of surprise in creative thinking first put forward by the pragmatist philosopher Peirce (Gonzalez & Haselager, 2005). Munnich et al. (2019) reflected this shift from simple perceptual schema disruption to sense-making when they suggest that there are two different ways the term "to be surprised" is commonly used—the first relating to an event which was not expected and the second relating to an event which cannot be explained. To draw from Boden (2004, p. 42) again,

> here this type of creativity is concerned, we have to do with expectations not about probabilities, but about possibilities. In such cases, our surprise at the creative idea recognizes that the world has turned out differently not just from the way we thought it would, but even from the way we thought it could.

In other words, existing schemas cannot incorporate what is seen by the human agent, so these need to be updated. Sense-making theories argue that surprise has value less because of its association with schema disruption and more because of the cognitive processes required to find an explanation for the event's occurrence (Foster & Keane, 2019; Munnich et al., 2019). From this perspective, simple perceptual disruption would not be enough to generate downstream benefits.

Alongside these cognitive explanations is a third reason why experiencing surprise may increase originality: Its relationship with affect. In contrast to the cognitive explanations of schema disruption or sensemaking, affective surprise pertains to the feeling that arises upon encountering something unexpected. The valence of affective surprise, whether it is perceived as positive, negative, or neutral, is a subject of ongoing debate within the academic community (Neta & Kim, 2023; Ortony, 2022; Ortony & Russell, 2024). However, it is clear that surprise can be emotionally arousing (Noordewier et al., 2016). It is possible that neither expectancy disruption nor sensemaking causes an increase in originality but rather the feelings that follow the momentary experience of surprise. While the evidence is inconclusive, there is an overall positive relationship between positive affect and originality (Baas et al., 2008), especially between the forms of activating motivation that are likely to be associated with surprise (Sundquist & Lubart, 2022). Indeed, Filipowicz (2006) found that surprise was a mediator of the relationship between originality and positive affect when participants were primed with humorous videos prior to undertaking a creative task, suggesting that the positive experience associated with surprise may explain any increase in originality found after a surprising event.

Therefore, there are three theoretically robust explanations as to why encountering surprising information in the form of an unexpected intrusion should increase creative ideation: (a) it disrupts existing schemas and directly seeds original ideas; (b) it orients people to creating meaning out of what they have seen and, in so doing, new thoughts are generated; and (c) it generates a feeling of positive and arousing affect. All three predict a positive relationship between encountering a surprising event and generating original thoughts.

While it is often assumed that surprise feeds creativity, and there are plausible theoretical accounts of the relationship between surprise and creativity, direct empirical evidence for the role of surprise in creativity remains scarce. This makes it hard to adjudicate between the different positions. Where empirical research exists, it is often intertwined with other factors, making it difficult to isolate its true effects. This isolation is important if we are to model the specific role of surprise in creative thinking or understand which of the three mechanisms is most important. For example, there is a rich literature on the relationship between diversifying experiences and creativity (Damian, 2025). Much like a surprise, a diversifying experience disrupts people's expectations by exposing them to something that modifies their existing models and generates new categories, resulting in more creative thinking. For example, when people were invited to an underwater disco, they were more likely than controls to generate original ideas in a subsequent divergent thinking task (Richardson et al., 2022). Similarly, manipulated "impossible" experiences are closely related to originality and flexibility in subsequent creativity tasks (Wiseman & Watt, 2022). As an illustration of this, Ritter et al. (2012) exposed participants to unexpected schematic violations which violated laws of physics (a suitcase that decreased in size as you approached it, a perception of walking faster than possible and witnessing a bottle flying in the air rather than off a table). This led to increased flexibility in the categories of responses to a follow-up unrelated Alternative Uses Task, although the originality of the ideas was not reported.

However, diversifying or impossible experiences capture something more complex than just a feeling of surprise. Such experiences are more likely to be emotionally arousing and include powerful emotions such as awe (Chirico et al., 2018, 2020), and they are more likely to require sensemaking activities to update prior understanding. In addition, they do not represent a sharp yet transient disruption in the same way that a surprising event does. So, even if surprise may be one aspect of their effects, they do not allow those effects to be clearly isolated. In order to fully understand the relationship between these two complex phenomena (surprise and creativity), we need a more precise approach that can allow us to explore separately the explanatory mechanisms. For instance, to test the assumption that schema disruption is important, we need

to isolate the proposed mechanisms for investigation. This is the main research question that drives this article.

Although Ritter et al. (2012) found increased flexibility of responses when violating the expected order of a typical process (preparing a traditional Dutch breakfast), other research has found that such small-scale schema discrepancy did not affect subsequent creativity. For example, research on encountering something random (in this case, Wikipedia pages), which required no sensemaking and had no affective component prior to being asked to take part in a divergent thinking task, did not increase the originality of the subsequent ideas (Malthouse et al., 2022). In addition, not all people use seemingly irrelevant information from the environment to inspire ideas in divergent thinking tasks or as hints in creative problem solving (Agnoli et al., 2015; Chronicle et al., 2001; Ross & Vallée-Tourangeau, 2022). This suggests that the relationship between experiencing surprise and subsequent creative thinking might be less straightforward than expected.

Surprise and Divergent Thinking

Divergent thinking is a core process in creative ideation. In divergent thinking tasks, participants are invited to generate several creative different ideas for the same object. If experiencing surprise seeds novel ideas, then it should be reflected in the originality of the generated ideas. This hypothesis is most commonly tested using the Alternative Uses Task (AUT; Guildford & Hoepfner, 1971), a test of divergent thinking that requires participants to generate alternative and original uses for an everyday object. In the studies reviewed above, each measure of originality was taken for after participants were primed with an experience that disrupted their existing schema in one area, however they were asked to generate uses for an object in another domain. Thus, results from these studies may be caused by the ongoing process of sensemaking or by an affective response rather than the direct seeding of ideas through perceptual disruption.

To the best of our knowledge, no research has tested the core prediction of the current study: Experiencing surprise about an object should increase the subsequent originality of uses involving that object. To isolate the effect of schema disruption, we presented incongruent objects in visually familiar scenes, then assessed whether this local violation sparks more original ideas about these same objects. This task allows us to investigate if simple schema disruption is enough to break fixation and seed a novel thought or whether the relationship between surprise and a new thought is more complicated than this. The simplicity of the task supports its role as a method for eliciting surprise, which is divorced from other potential mechanisms.

The Current Study

We report four preregistered studies (total $N = 964$) that sought to investigate empirically the theoretical link between creative cognition and surprise, operationalized as schema disruption. We specifically focused on the surprise associated with manipulations of the semantic congruency of an object, that is, how much it belongs to a scene (Öhlschläger & Võ, 2017) to isolate the effects of schema disruption and investigate its role in ideation. We subsequently invited participants to take part in a traditional divergent thinking task (the AUT; Guildford & Hoepfner, 1971; Reiter-Palmon et al., 2019). This paradigm assesses idea generation by asking participants to generate as many ideas as they can for an everyday object and then scoring those uses for originality. Thus, we could investigate the direct effect of schema disruption on original ideas for the surprising object postdisruption.

We progressively tested a series of preregistered hypotheses. These hypotheses initially assumed a positive relationship between surprise and originality in line with the predictions of a schema disruption explanation of surprise. Experiment 1a used images showing an object both in and out of place and measured both how surprising participants found the object and the subsequent originality of the uses generated for the object. Experiment 1b was an internal replication of this study. Experiment 2 replicated Experiment 1 with a different set of stimuli, which investigated the effect of a surprising use for the object rather than just its placement. Experiment 3 replicates the findings of Experiment 2 but also ensured to control for both the priming effect of order of conditions and the believability of the picture. Through these four experiments, we explored in experimentally controlled conditions whether the surprise associated with different forms of norm violations can affect creative performance.

Experiment 1a

Method

Participants

A power calculation for the main hypothesis of interest—that is, that originality score can be predicted by the rated surprise—was conducted using GLIMMPSE (Kreidler et al., 2013), suggesting a total sample size of 90 participants to detect a small effect with a power of 0.82. Given the likely attrition rate, 120 participants were recruited from Prolific.co over three linked sessions. They were paid an overall fee of £9 on completion of Session 3. One hundred and seventeen participants completed all three sessions and were included in the analysis; 58 were men with an average age of 41.31 years ($SD = 12.51$), 55 were women with an average age of 40.15 ($SD = 11.19$), three preferred not to say ($M_{age} = 45.00$, $SD = 11.19$) and one was nonbinary (age = 37).

Materials

Experiment Stimuli. The experimental trials consisted of images taken from the Scene Grammar (SCEGRAM; Öhlschläger & Võ, 2017) database of images of everyday objects placed in different levels of congruity with their environment. Öhlschläger and Võ suggest that there is a logic that guides our expectations of how to view a scene. They call this logic, a grammar. This grammar consists of both semantic and syntactic expectations, with the addition of possible and impossible positions. For example, in the control condition (CON), a mug is shown in the kitchen and in the traditionally expected position (e.g., on the kitchen sideboard). In the semantic violation (SEM) condition, the object is incongruent for the setting (e.g., an umbrella in a bathroom), but it is in a position that makes sense (e.g., folded next to the toilet; a semantic violation). In the syntactic violation (SYN) condition, the object belongs in the room but is put in a location, which is unexpected but physically possible (e.g., the mug on a chair and a syntactic violation). In the semantic and syntactic violation (SEMSYN) condition, the object breaks both syntactic and semantic expectations (e.g., an umbrella

on the toilet in the bathroom). In the extreme syntactic violation (EXSYN) condition, a semantically consistent object experiences an extreme syntactic dislocation, in other words, in an impossible position but in the expected room (e.g., a mug floating in midair in the kitchen). Finally, the extreme semantic and syntactic violation (EXSEMSYN) condition breaks the semantic expectations and presents an extreme syntactic dislocation (e.g., a bicycle helmet floating in the air in the kitchen).

We selected this stimulus set (with the six conditions) as it allowed us to investigate schema violation in a graduated manner. Our primary interest was in the effect of seeing novel object–scene combination, and because the effect may be affected by the stimulus variation, we explicitly addressed that in the analyses and stimulus selection.

Measures of Originality. The outcome variable for this experiment was the level of originality of the responses to an AUT. For each trial, the image of the object[1] was presented embedded in the scene (see Figure 1), but the image was removed when the participant was invited to generate ideas for the object. The participants were given 90 s to complete the task using the prompt,

> Can you come up with original and creative uses for the [OBJECT]. You can enter as many ideas as you like in 90 SECONDS. Creative quality is more important than quantity. It's better to have a few really good ideas than a lot of uncreative ones.

They were prompted to press [ENTER] after each idea. The responses were subsequently hand checked by Wendy Ross, and those responses separated by a comma, full stop, or semicolon were treated as individual responses.

We preregistered two ways of assessing originality. First, we assessed semantic distance between the object and the response using the SemDis (Beaty & Johnson, 2020) open scoring platform. We also assessed the originality ideas by using the Open Creativity Scoring with AI (OCSAI; Organisciak et al., 2023), which has a correlation of .8 with human raters. Both SemDis and OCSAI gave the same outcomes, and so, for reasons of space, we report only the OCSAI. All other results can be found as well as the additional online materials at https://osf.io/4n8tc/?view_only=5f82eebeeac347b6a084f28a0356109a.[2]

Design and Procedure

The experiment was presented online using PsychoPy (Peirce et al., 2019) for the experimental trials, and participants were redirected to Qualtrics for the measures of individual differences. Demographic information collection and condition allocation were conducted using Vespr[3] , which works with PsychoPy's online version (Pavlovia) to counterbalance groups. The trials were presented in three sessions of six trials (one for each condition), followed by the measures of individual differences. Participants were told on recruitment that they must complete all three sessions to receive their reward. They were also told to leave a minimum of 1 hr and a maximum of 24 hr between sessions. This was to avoid participant fatigue.

Each trial consisted of an image taken from the SCEGRAM database. Participants were prompted to locate the object and press any key on location (the time spent for the location was recorded). Once they had pressed a key, the following three questions appeared: (a) How surprising do you find the object's location (b) How interesting do you find the object in that location object? (c) How useful do you find the object in this location? These questions were answered on a slider from 1 to 10, with 0 marked as *not at all* and 10 marked as *very* (see Figure 1).

On the following screen, participants were asked to come up with original and creative uses for the object in 90 s. A timer in the top right-hand corner displayed the time to the end of the trial. At the end of six trials, participants were directed to Qualtrics for the measures of individual differences.

The design was a fully within-participants design with each participant experiencing all 18 trials. Each participant saw 18 different objects across three blocks, and the objects varied systematically across six dimensions. The object and their coherence were counterbalanced, and no participant saw the same object more than once.

Preregistered Hypotheses

The preregistration of these hypotheses can be found here at https://osf.io/qwxb5/?view_only=6da4cc13dcd64901b8740bc62111fa3a.

Hypothesis 1 (H1): Subjective ratings of surprise will increase as the congruency of the object decreases. We predict CON < SEM < SYN < SEMSYN < EXSYN < EXSEMSYN.

Hypothesis 2 (H2): Subjective ratings of usefulness will decrease as the congruency of the object decreases. We predict CON > SEM > SYN > SEMSYN > EXSYN > EXSEMSYN.

Hypothesis 3 (H3): The more surprising the object, the longer it will take to locate.

Hypothesis 4 (H4): Ideas generated for the object will be more creative the more surprising the object placement.

Hypothesis 5 (H5): Ideas generated for the object will be less creative the more useful the object appears.

Hypothesis 6 (H6): Ideas generated for the object will be more creative the longer it takes to find it.

Results

All statistical analyses were conducted using R. The analysis was conducted using lme4 (Bates et al., 2015) at the level of single response to the AUT task with participant, stimuli, and block coded as random factors and trial number nested within participant and block.[4] To better understand the relationship between

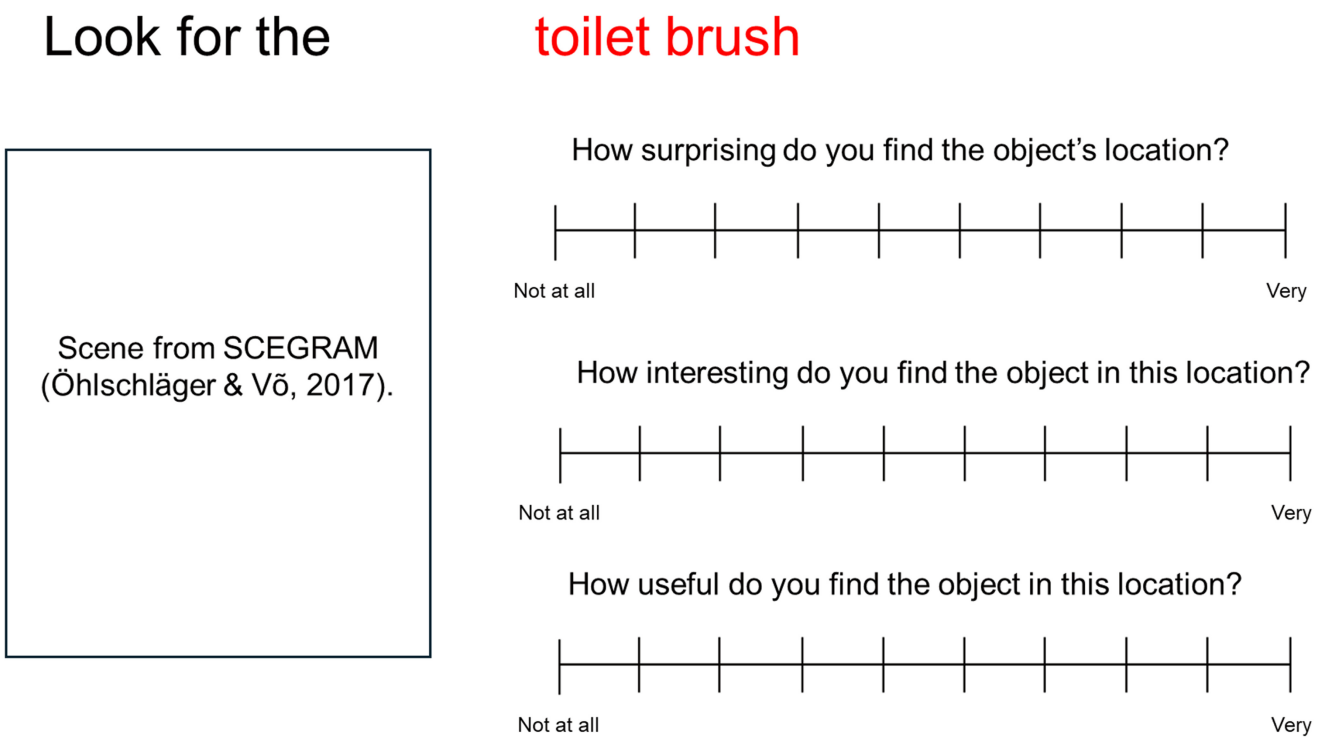
[1] Objects were taken from this list: lamp, bowl, hanger, glove, soap, computer mouse, cake pan, toilet brush, strainer, egg carton, paper moulds, toothpaste, card game, CD, plate, cheese grater, phone, oven mitt, vase, pot lid, toast, watercolour, rolling pin, washing up liquid, clock, umbrella, DVD cover, hat, helmet, hairband, and dish soap.

[2] Measures of individual differences: We preregistered exploratory analysis of four different individual difference scales (the Interpersonal Reactivity Index, Davis, 1983; the Need for Cognition Short Scale 6, Lins de Holanda Coelho et al., 2020; the Revised Personality Inventory Personality Scale, Costa & McCrae, 2008; and the Sussex Cognitive Styles, Mealor et al., 2016), which we theorized may have a relationship with originality and surprise. There were no significant relationships between these and measures of surprise or originality (SEMDIS or OCSAI) and the results of this analysis can be found at https://osf.io/4n8tc/?view_only=5f82eebeeac347b6a084f28a0356109a.

[3] https://moryscarter.com/vespr/portal/.

[4] The random effect structure was coded in lmer using the following syntax: (1|ID) + (1|item) + (1|ID/trialN) + (1|expt) + (1|expt/trialN).

Figure 1
Object Presentation and Self-Report Scales



Note. SCEGRAM = Scene Grammar. See the online article for the color version of this figure.

categorical predictors and outcomes, the mixed models were converted to either analyses of variance (ANOVA) or deviance.

Exclusions

Upon visual examination of the scatter plot, a number of outliers for response time were identified, which led to spurious correlations. The range of response times was from less than 1 to 1,616.08 s. A decision was made to exclude all trials where the response time was over 120 s. This led to the exclusion of 295 responses (4% of the total) and left 7,118 uses for analysis. The analysis of the full data set without exclusions can be found at https://osf.io/4n8tc/?view_only=5f82eebeeac347b6a084f28a0356109a.

Experimental Manipulation

Self-Reports of Reactions to the Objects. We first assessed whether our experimental manipulation had worked, and the different object placements had evoked different responses in our participants. The average responses across experimental conditions on the three items—surprisingness, usefulness, and interestingness—are reported in Table 1.

H1: Subjective Ratings of Surprise Will Increase as the Congruency of the Object Decreases: We Predicted That CON < SEM < SYN < SEMSYN < EXSYN < EXSEMSYN. The ANOVA suggests that the main effect of experimental condition is statistically significant and large, $F(5) = 996.39$, $p < .001$, $\eta_p^2 = .48$, 95% confidence interval (CI) [0.46, 1.00]. Post hoc tests show that our hypothesis was not fully supported; there were significant differences between all conditions ($p_{Bonf} < .001$) except between SEM and SEMSYN ($p_{Bonf} = .083$). The direction of the change was different from that hypothesized—semantic violations were rated as more surprising than syntactic violations. The final order of feelings of surprise was CON < SYN < EXSYN < SEM < SEMSYN < EXSEMSYN.

H2: Subjective Ratings of Usefulness Will Decrease as the Congruency of the Object Decreases: We Predict CON > SEM > SYN > SEMSYN > EXSYN > EXSEMSYN. The ANOVA suggests that the main effect of condition is statistically significant and large, $F(5) = 856.38$, $p < .001$, $\eta_p^2 = .41$, 95% CI [0.39, 1.00]. Post hoc tests show that our hypothesis was not fully supported, and the direction of effect was not as predicted. There were significant differences between all conditions (highest $p_{Bonf} = .004$) except between EXSYN and SYN ($p_{Bonf} > .999$). Syntactic violations appeared to still be seen as relatively useful. The final order of ratings of usefulness was CON > SYN > EXSYN > SEM > SEMSYN > EXSEMSYN.

Interestingness. We also measured interestingness, but we preregistered no particular prediction. The ANOVA suggests that the main effect of condition is statistically significant and large,

Table 1
Average (Standard Deviation) Response for Usefulness, Surprise, and Interestingness as a Function of Condition

Measure	CON	SEM	SYN	SEMSYN	EXSYN	EXSEMSYN
Surprising	1.65 (1.41)	7.41 (2.34)	4.32 (2.87)	7.67 (2.41)	6.06 (3.10)	8.80 (1.67)
Usefulness	8.76 (1.75)	2.91 (2.09)	5.92 (2.78)	2.53 (1.76)	5.77 (3.12)	2.23 (1.67)
Interesting	2.29 (1.63)	4.69 (2.58)	3.44 (2.28)	5.08 (2.69)	4.75 (2.87)	5.51 (2.92)

Note. Values are expressed as mean (standard deviation). CON = control condition; SEM = semantic violation condition; SYN = syntactic violation condition; SEMSYN = semantic and syntactic violation condition; EXSYN = extreme syntactic violation condition; EXSEMSYN = extreme semantic and syntactic violation condition.

$F(5) = 293.93$, $p < .001$, $\eta_p^2 = .34$, 95% CI [0.32, 1.00]. Post hoc tests showed that there were significant differences between all conditions (highest $p_{Bonf} = .047$) except between EXSYN and SEM ($p_{Bonf} > .999$). The final order of how interesting the object was as CON < SYN < SEM < EXSYN < SEMSYN < EXSEMSYN.

Behavioral Measures. We checked whether participant surprise could be triangulated with objective behavioral measures, namely, latency to find the object. The relationship is illustrated in Figure 2. We hypothesized that the longer it took to notice an object, the more surprising it would be rated.

H3: The More Surprising the Object, the Longer It Will Take to Locate. We fitted a linear mixed model to predict surprise with response time. Within this model, response time is a statistically nonsignificant and negative predictor of surprise, $b = -0.01$, 95% CI [−0.02, <0.001], $t(7240) = -3.10$, $p = .002$, $\beta = -.03$, 95% CI [−0.05, −0.01]. Our hypothesis was not supported—the significant effect was in the opposite direction. However, on inspection of the scatterplot in Figure 2, we suggest that this result should be treated with caution given the heteroscedasticity of the data.

Relationship With Originality

We next assessed the differences in originality across the six groups using OCSAI (see Table 2).

The ANOVA suggests that the main effect of condition is statistically not significant and very small, $F(5) = 0.97$, $p = .433$, $\eta_p^2 < .01$, 95% CI [0.00, 1.00].

H4: Ideas Generated for the Object Will Be More Creative the More Surprising the Object Placement. The relationship between levels of surprise and originality is illustrated in Panel A of Figure 3. We fitted a linear mixed model, which suggested that the surprise is a statistically nonsignificant predictor of originality, marginal $R^2 < .001$, $b = <0.001$, 95% CI [0.00, 0.01], $t(7240) = 0.70$, $p = .486$, $\beta = <.001$, 95% CI [−0.02, 0.03].

H5: Ideas Generated for the Object Will Be Less Creative the More Useful the Object Appears. The relationship between levels of perceived usefulness and originality is illustrated in Panel B of Figure 3. Linear mixed models were used to examine the effects of perceived usefulness on originality and semantic distance. The results indicate that perceived usefulness is a statistically nonsignificant predictor of originality, marginal $R^2 = <.001$, $b = <0.001$, 95% CI [−0.01, 0.01], $t(7240) = 0.08$, $p = .935$, $\beta = <.001$, 95% CI [−0.03, 0.03]).

Interestingness. We also measured the interestingness of the object, but we did not preregister a hypothesis. The relationship is illustrated in Panel C of Figure 3. Linear mixed models suggest that perceived interestingness is a statistically nonsignificant predictor of originality, marginal $R^2 < .001$, $b < -0.001$, 95% CI [−0.009, 0.005], $t(7240) = -0.57$, $p = .569$, $\beta < -.001$, 95% CI [−0.03, 0.02].

Discussion

The data show that we successfully manipulated levels of the reported surprisingness of an object across experimental conditions. The level of incongruency was related to the surprise ratings, albeit not always in the predicted direction. Despite the success of these manipulations, there was no difference in the originality between conditions, and surprise was not significantly related to originality across the conditions. In addition, neither the perceived usefulness of the object nor the interest it elicited was significantly related to originality. It seems likely, therefore, that simply manipulating the participants' expectancies did not lead to increased originality of response.

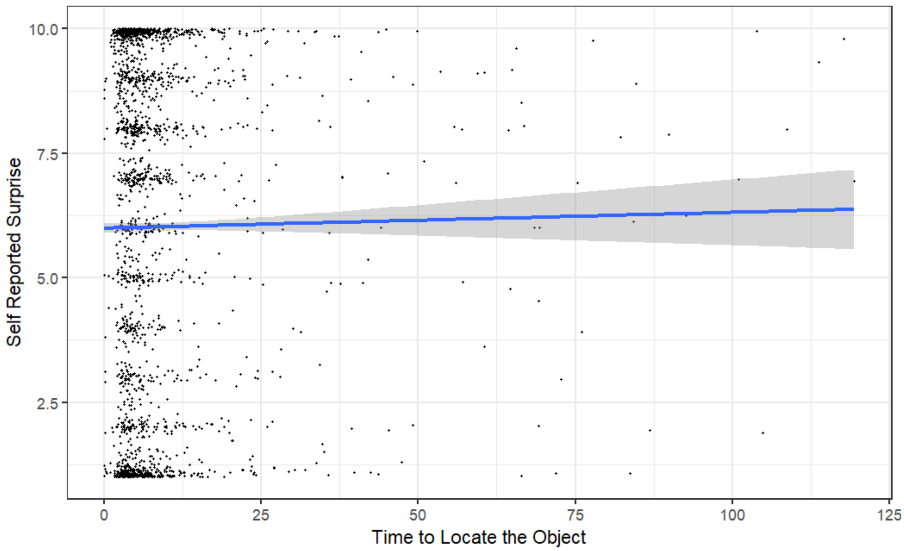
The wide range of response times, however, indicates a potential methodological issue. Participants were invited to press a button when they saw the target object, but there was no way of verifying that they had found the object. In addition, if the measure was a true measure of searching time, then a searching time of 120 s is likely to have led to cognitive fatigue. In addition, 18 similarly structured trials are likely to lead to boredom, although we included trials as nested factors in our analysis; a reduction in trials would ensure that the null effect was not down to cognitive fatigue. Finally, each response was hand checked by Wendy Ross, and on reading them, it was observed that many were anchored within the scene pictured. There is some evidence that presenting objects in scenes reduces overall originality levels (George et al., 2023) compared to presenting simply the word, and this may be contributing to a floor effect across conditions. To mitigate for this last issue, we added in a further study the following line to the instructions: "The uses can be any you can think of. You do not have to think of uses for the object in the location pictured. Let your imagination run wild!"

Experiment 1b

In Experiment 1b, we aimed to replicate the findings from Experiment 1a. To ensure that participants were not fatigued, we reduced the number of conditions down to two—the congruent (CON) and the extremely incongruent (EXSEMSYN) conditions, and therefore could replicate the number of trials (three in each condition) in one block. Second, we refined the procedure so that participants were required to click on the object rather than press a button

Figure 2

The Relationship Between Surprise and Time Taken to Locate the Object



Note. See the online article for the color version of this figure.

when they found it. Third, we introduced a button to allow participants to skip a trial after 20 s when they had not been able to find the object to better account for cognitive load and to ensure that response times are accurate.

Method

Participants

We recruited 150 participants from Prolific.co, and data were saved for 143 participants; 73 participants were men ($M_{age} = 37.19$, $SD = 13.28$), 68 were women ($M_{age} = 38.01$, $SD = 14.47$), and two preferred not to say ($M_{age} = 37$, $SD = 4.24$).

Procedure and Design

The procedure was identical to Experiment 1a except for the differences outlined above: (a) a reduction in conditions and trials, (b) requiring participants to select the object when they found it, (c) the additional prompt to think beyond the picture, and (d) the chance to move on after 20 s.

Preregistered Hypotheses

The preregistration for these hypotheses can be found at https://osf.io/gpsrb/?view_only=a90fa405a8a24044be9d68b426774a68.

Hypothesis 1 (H1): Participants will be more surprised in the extremely incongruent condition than in the congruent condition.

Hypothesis 2 (H2): Originality as measured by OCSAI will not be predicted by surprise, usefulness, or interestingness.

Results and Discussion

In line with the preregistration, we excluded all trials where the response time was over 30 s. This removed 14 ideas from the data set, leaving us with 2,930 data points.

H1: Participants Will Be More Surprised in the Extreme Condition

We replicated the findings from Experiment 1. Self-reported ratings of surprise were higher in the extreme condition ($M = 8.52$, $SD = 1.72$) than in the control condition ($M = 1.72$, $SD = 1.23$). This difference was significant, $F(1) = 3{,}401.79$, $p < .001$, $\eta_p^2 = 1.00$, 95% CI [0.99, 1.00].

H2: Originality as Measured by OCSAI Will Not Be Predicted by Surprise, Usefulness, or Interestingness

We replicated the findings from Experiment 1. Originality was broadly similar across both conditions. In the congruent condition,

Table 2

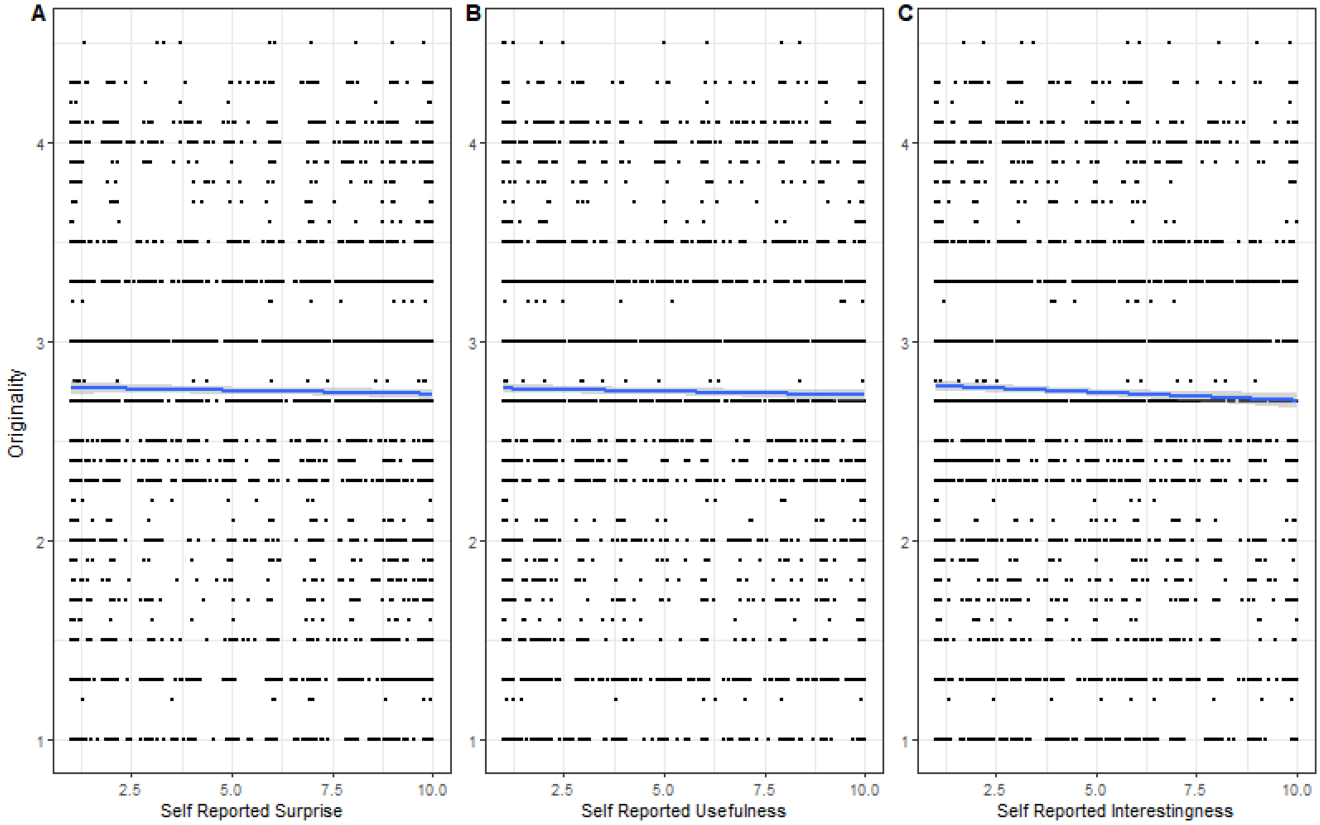
Average (Standard Deviation) Originality of Items as a Function of Condition

Measure	CON	SEM	SYN	SEMSYN	EXSYN	EXSEMSYN
Originality	2.74 (0.72)	2.77 (0.72)	2.74 (0.75)	2.76 (0.72)	2.76 (0.72)	2.72 (0.74)

Note. Values are expressed as mean (standard deviation). CON = control condition; SEM = semantic violation condition; SYN = syntactic violation condition; SEMSYN = semantic and syntactic violation condition; EXSYN = extreme syntactic violation condition; EXSEMSYN = extreme semantic and syntactic violation condition.

Figure 3

Relationship Between Originality and the Outcome Measures



Note. Relationship between originality and (A) surprise, (B) usefulness, and (C) interestingness. See the online article for the color version of this figure.

it averaged 2.65 ($SD = 0.76$), while in the extreme condition, it averaged 2.66 ($SD = 0.77$). The ANOVA suggests that the main effect of condition is statistically not significant, $F(1) = 0.34$, $p = .558$, $\eta_p^2 < .001$, 95% CI [0.00, 1.00].

We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict originality with surprise, interestingness, and usefulness. The model's explanatory power related to the fixed effects alone (marginal R^2) is <.001. Surprise is a statistically nonsignificant predictor of originality, $b < 0.001$, 95% CI [−0.02, 0.01], $t(2921) = -0.55$, $p = .582$, $\beta = -.02$, 95% CI [−0.10, 0.05]. Interestingness is a statistically nonsignificant predictor of originality, $b < 0.001$, 95% CI [−0.01 0.02], $t(2921) = 0.71$, $p = .476$, $\beta = .02$, 95% CI [−0.03, 0.07]. Usefulness is a statistically nonsignificant predictor of originality, $b < 0.001$, 95% CI [−0.02, 0.01], $t(2921) = -0.76$, $p = .448$, $\beta = -.02$, 95% CI [−0.09, 0.04].

We therefore replicated the main effects of Experiment 1a.

Experiment 2

The findings from Experiment 1a suggest that while our experimental manipulation worked—we successfully elicited a feeling of surprise in participants—this did not affect originality levels. We replicated this effect in Experiment 1b giving us some confidence that presenting participants with objects in a situation that they find surprising is not enough to inspire more creative ideas about those objects.

The scenes in Experiment 1 were static scenes with the object out of place. The surprise elicited in Experiment 1 came from an incongruency in the placement of the object, but incongruencies can also arise from the unusual manipulation of an object by a human. In Experiment 2, we aimed to further investigate the unexpected findings that simple perceptual surprise does not support creative ideation on a similar stimulus set—the ObjAct stimulus set (Shir et al., 2021). The ObjAct stimulus set presents a series of still pictures where humans are manipulating either a congruent or an incongruent object. For example, a woman can be seen either playing 10-pin bowls with a bowling ball or a cabbage or a child can be seen eating an ice cream or a light bulb.

In addition, it may be that people will rate objects as less surprising after several trials, because they become habituated, dampening down the effect of surprise on originality, and so we also preregistered an analysis of the hypotheses in (a) the overall data set and (b) just on the first trial.

Method

Participants

We recruited 320 participants to take part. Three hundred sixteen participants completed the study. One hundred fifty-eight were men with an average age of 40.83 ($SD = 14.97$), 155 were women with an

average age of 42.57 ($SD = 13.80$), one was nonbinary (age = 22), one preferred not to say (age = 48), and one identified another way (age = 50). We recruited a larger number of participants than in Experiments 1a and 1b to allow for between-subject analysis of the first trial as in the preregistration.

Materials and Measures

We selected three congruent and three matching incongruent stimuli (six stimuli in total) at random from the ObjACT database.

Procedure and Design

The procedure was identical to Experiment 1 except for the difference in stimulus set. All participants took part in all six trials.

Preregistered Hypotheses

The preregistration for these hypotheses can be found at https://osf.io/h67yx/?view_only=2681d7fae97a4679aee78de44bcd1bcf.

Hypothesis 1 (H1): Participants will be more surprised in the incongruent condition.

Hypothesis 2 (H2): Originality as measured by OCSAI will not be predicted by surprise, usefulness, or interestingness.

Results

H1: Participants Will Be More Surprised in the Incongruent Condition

Full Set of Trials. The average reported surprise was 1.46 ($SD = 1.25$) in the congruent condition, whereas it was 8.52 ($SD = 2.02$) in the incongruent condition. The ANOVA suggests that the main effect of condition is statistically significant and large, $F(1) = 322.23$, $p < .001$, $\eta_p^2 = .97$, 95% CI [0.92, 1.00].

First Trial. Again, for the first trial, in the congruent condition, surprise was 1.27 ($SD = 0.67$), while in the incongruent condition, it was 7.73 ($SD = 2.49$). The ANOVA suggests that the main effect of condition is statistically significant and large, $F(1) = 2{,}426.59$, $p < .001$, $\eta_p^2 = 0.71$, 95% CI [0.69, 1.00].

H2: Originality as Measured by OCSAI Will Not Be Predicted by Surprise, Usefulness, or Interestingness

Full Set of Trials. We next assessed if originality as measured by OCSAI differed between the two conditions. Originality in the congruent condition was lower ($M = 2.47$, $SD = 0.69$) than in the incongruent condition ($M = 2.68$, $SD = 0.70$). This difference is statistically not significant and small, $F(1) = 0.39$, $p = .548$, $\eta_p^2 = .04$, 95% CI [0.00, 1.00].

We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict originality with surprise, usefulness, and interestingness. The model included ID, item, and trial number as random effects.[5] Within this model, surprise is a statistically nonsignificant predictor of originality, $b < 0.001$, 95% CI [0.00, 0.02], $t(7133) = 1.23$, $p = .218$, $\beta = .04$, 95% CI [−0.02, 0.10], interestingness is a statistically nonsignificant predictor of originality, $\beta = -.002$, 95% CI [−0.01, 0.01], $t(7133) = -0.53$, $p = .596$; $\beta < -.001$, 95% CI [−0.04, 0.02]. However, usefulness is a statistically significant predictor of originality, $b = 0.01$, 95% CI [0.00, 0.02], $t(7133) = 2.38$, $p = .017$, $\beta = .05$, 95% CI [0.00, 0.09]. Given the other findings, we suggest that this last significant relationship should be treated with caution.

First Trial. We fitted a linear mixed model on the first trial; within this model surprise is a statistically nonsignificant predictor of originality, $b = 4.29e{-}03$, 95% CI [−0.02, 0.03], $t(1395) = 0.31$, $p = .754$, $\beta = .02$, 95% CI [−0.11, 0.16], interestingness is a statistically nonsignificant predictor of originality, $b = -0.01$, 95% CI [−0.04, 0.01], $t(1395) = -0.91$, $p = .363$, $\beta = -.04$, 95% CI [−0.14, 0.05], and usefulness is a statistically nonsignificant predictor of originality, $b = -0.004$, 95% CI [−0.02, 0.02], $t(1395) = -0.41$, $p = .680$, $\beta = -.02$, 95% CI [−0.12, 0.08].

Discussion

As in Experiment 1, while the different conditions elicited reliably different levels of surprise, they did not differ in terms of the originality of uses that participants could generate. This means that even when the objects are in motion—when someone is doing something with them—surprising placements do not act as an inspiration.

The analysis of the first trial showed a lower reported surprise in the first trial than in subsequent ones, which was unexpected. Since the presentation order of the stimuli was randomized, we could not exclude that this result could be influenced by the type of stimulus seen at the first trial (incongruent vs. congruent). However, we could hypothesize that the participants who experience a congruent first trial could be more surprised when they then experienced an incongruent one, which would act as a contrast class. We therefore designed Experiment 3 to test this hypothesis.

Experiment 3

Experiment 2 supported the findings from Experiment 1 that higher levels of surprise emerging from schema violation did not elicit originality. When we analyzed the data from the first trial, it seemed counterintuitive that the surprise levels overall were lower. This could be a potential confound, so we designed the final experiment to take account of any possible priming effects. In addition, it is possible for a participant to be surprised without this eliciting a clear schema violation about the object—therefore not "true" surprise. This is more likely with images that can be surprising without being believable. We therefore assessed whether participants believed in the stimulus rather than assuming it was either photoshopped or generated through methods of artificial intelligence.

Method

We aimed to replicate the findings using the same stimulus set as Experiment 2. The experimental procedure was broadly similar aside from the following changes. To fully reduce fatigue, only two trials were presented, and these were divided into four orders: congruent followed by incongruent, congruent followed by incongruent, incongruent followed by congruent and incongruent followed by incongruent. This allowed us to assess if surprisingness was higher following a

[5] Formula: (∼1 | prolific, ∼1 | item, ∼1 | trials.thisN:prolific, ∼1 | prolific).

control condition. We also asked participants if they thought the photo depicted a real event prior to asking for creative ideas.

Participants

We again recruited to be able to assess the between-subject effect of order. In addition, recruitment for this study was also tied to recruitment for a separate, unrelated study. We therefore recruited 400 participants, and data were retained for 388. There were 194 men ($M_{age} = 40.71$, $SD = 13.45$) and 189 women ($M_{age} = 40.90$, $SD = 13.69$). Three people identified as nonbinary ($M_{age} = 31.33$, $SD = 5.69$), one identified another way (aged 65), and one preferred not to disclose (aged 28).

Materials

We used the same stimulus set as Experiment 2, but rather than random selection of images, two members of the research team rated each of the incongruent stimuli for believability, operationalized as how much they believed the image was real. The final stimuli were randomly selected from those deemed most believable by the research team and consisted of a small child in a crowd eating an ice cream (congruent) and a lightbulb (incongruent) and a different child on a beach digging next to a bucket (congruent) and a large cake (incongruent)

Preregistered Hypotheses

The preregistration for these hypotheses can be found at https://osf.io/2d4h9/?view_only=6558258720654317bc45fd225743ccc3.

Hypothesis 1 (H1): There is a priming effect of surprise so that participants primed on a nonsurprising image will experience higher levels of surprise than those primed on a control condition.

Hypothesis 2 (H2): A believable image will generate higher levels of originality when featured in a follow-up AUT task.

Results

Again, we analyzed the results at the level of individual answers to the AUT. We had 1,783 valid answers in the congruent conditions and 1,628 answers in the incongruent condition.

Manipulation Check

The average ratings in each condition are presented in Table 3. An ANOVA suggests that for surprise, the main effect of condition is statistically significant and large, $F(1) = 429.36$, $p = .002$, $\eta_p^2 = 1.00$, 95% CI [0.92, 1.00]. For interestingness, the main effect of condition is statistically significant and large, $F(1) = 337.50$, $p = .002$, $\eta_p^2 = .99$, 95% CI [0.90, 1.00], and for usefulness the main effect of condition is also statistically significant and large, $F(1) = 429.36$, $p = .002$, $\eta_p^2 = 1.00$, 95% CI [0.92, 1.00]. We are therefore confident that, as in previous experiments, a robust difference between our conditions has been found.

Table 3
Average (Standard Deviation) Response for Usefulness, Surprise, and Interestingness as a Function of Condition

Measure	Congruent	Incongruent
Usefulness	6.87 (2.89)	2.46 (1.92)
Surprisingness	1.72 (1.69)	8.16 (2.41)
Interestingness	2.86 (1.95)	6.12 (2.50)

Order Effects on Surprise

H1: There Is a Priming Effect of Surprise so That Participants Primed on a Nonsurprising Image Will Experience Higher Levels of Surprise Than Those Primed on a Control Condition. We also tested for order effects on the surprisingness of the objects. To do this, we created four experimental groups: congruent to congruent, congruent to incongruent, incongruent to congruent, and incongruent to incongruent. At Time 2, surprise was higher in the two incongruent conditions than in the congruent conditions, as displayed in Table 4.

To test H1, we were particularly interested in the difference between the two incongruent conditions. If participants were primed to expect a neutral image after seeing a first, then we would expect an increase in surprise when the incongruent condition follows the congruent condition. The ANOVA suggests that the main effect of group is statistically not significant, $F(1) = 0.01$, $p = .914$, $\eta_p^2 < .001$, 95% CI [0.00, 1.00]. H1 was not supported. People were not more surprised if they had been primed on a congruent condition.

Relationship With Originality

Again, the average originality level of ideas was broadly similar across conditions. For the congruent condition, originality was slightly lower at 2.51 ($SD = 0.77$), and for the incongruent condition, it was 2.56 ($SD = 0.74$). The main effect of condition is statistically not significant, $F(1) = 1.16$, $p = .391$, $\eta_p^2 = .36$, 95% CI [0.00, 1.00].

We fitted a linear mixed model to assess whether surprisingness, interestingness, and usefulness predict originality. The model's total explanatory power is moderate (conditional $R^2 = .21$). Within this model, surprisingness is a statistically nonsignificant predictor of originality, $b = 0.004$, 95% CI [−0.010, 0.020], $t(3404) = 0.53$, $p = .597$, $\beta = .02$, 95% CI [−0.06, 0.10], interestingness is a statistically nonsignificant predictor of originality, $b = 0.002$, 95% CI [−0.01, 0.02], $t(3404) = 0.24$, $p = .807$, $\beta = .007$, 95% CI [−0.05, 0.06], and usefulness is a statistically nonsignificant predictor of originality, $b = -0.005$, 95% CI [−0.02, 0.01], $t(3404) = -0.70$, $p = .483$, $\beta = -.02$, 95% CI [−0.01, 0.04].

Believability

We also assessed in this experiment whether people found the image believable. As might be expected, 79% of people found the

Table 4
Average (Standard Deviation) Surprise as a Function of First Trial and Condition

	Time 2 condition	
Time 1 condition	Congruent	Incongruent
Congruent	1.50 (1.49)	8.42 (2.40)
Incongruent	1.59 (1.22)	8.38 (1.61)

congruent image believable, while only 14% of the incongruent condition did. The average originality of items for each of these groupings is shown in Table 5.

The ANOVA suggests that the main effect of condition is statistically not significant, $F(1) = 1.04$, $p = .409$, $\eta_p^2 = .33$, 95% CI [0.00, 1.00]. The main effect of belief is statistically not significant, $F(1) = 0.39$, $p = .531$, $\eta_p^2 < .001$, 95% CI [0.00, 1.00]; the interaction between condition and belief is statistically not significant, $F(1) = 1.99$, $p = .159$, $\eta_p^2 = 7.50e{-}04$, 95% CI [0.00, 1.00]. H2 was not upheld.

Discussion

We replicated the findings from Experiments 1 and 2. Participants' ratings of surprisingness of object placement, interestingness, and usefulness reliably differed between conditions. Again, this did not have an effect on the originality of ideas for what to do with the object. We also found no evidence of a priming effect, suggesting that participants' ratings of surprise remained stable across conditions. Finally, believing or not in the pictured use of the object did not seem to affect originality.

General Discussion

Over the course of four separate experiments, we manipulated participants' ratings of the surprisingness of the object placement, the usefulness of the object, and how interesting it was. Objects in unexpected placements were seen as more surprising than those placed in a way congruent with expectations. We can therefore be confident that this experimental manipulation is robust. However, this disruption of expectations had no downstream effect on the originality of the ideas for that object. Contrary to the folk conception of the relationship between surprise and creativity, it seems likely that simple schema disruption is not enough to generate originality of ideas. If there is a relationship between surprise and creative performance, then schema disruption cannot be the sole explanation.

Implications

There is a long folk association between surprise and being inspired to generate original ideas. Our data suggest that this claim needs to be treated with caution. We have established that simple perceptual schema disruption is not enough to see a significant increase in the originality of ideas for the disrupted object. We successfully manipulated whether participants found an object surprising across each of the four experiments. Moreover, this high level of surprise was also associated with increased interest in the objects. However, we did not see the hypothesized relationship between surprise and originality of ideas.

Arguably, the manipulation elicited a "perception" of surprise (i.e., a recognition that an unexpected event had occurred) rather than "a 'feeling' of surprise (i.e., an emotional reaction to that event)."[6] The persistence of the surprisingness ratings across several trials and the lack of an effect of priming suggest that, while participants could identify a surprising event, this could be disentangled from feeling surprised or from an increase of emotional arousal associated with the emotion of surprise. It seems unlikely that participants would continue to be surprised over the course of 16 trials. This suggests that current conceptions of surprise need to be updated: It is an open question whether one can assess a situation as surprising without feeling surprise or without being sufficiently activated by the situation to induce variations in the creative thinking processes. The results from this study suggest that cognitive surprise (recognizing something is surprising) should be treated as separate from affective surprise (the feeling of surprise).

Second, the current studies suggest that schema disruption is not enough to account for the increase in originality noted in theoretical, empirical, and folk accounts of the relationship between surprise and originality. Surprise does not work simply by seeding new ideas about what an object can do. We also tested in Experiment 3 how much participants were affected by the believability of the object placement. The data here suggest that even when an impossible situation is believed by a participant, it has no effect on their subsequent creativity, confirming that schema updating does not have an effect on the originality of ideas. Thus, while a perceptual disruption is necessary for surprise by definition, our data show that it is not sufficient to have a significant impact on the originality of ideas for that object. If there is a relationship between surprise and creativity, it appears that other aspects need to also be present. This has implications for theories of serendipity in scientific discovery (e.g., Ross, 2024) and interventions to increase creative thinking in both educational and organizational settings.

Limitations and Future Directions

There are other potential explanations for a relationship between surprise and creative ideation that were not investigated in this study. The experimental paradigm we used did not allow the participants to have agency over the surprising observation. Theories that discuss the importance of surprise to understanding human creativity emphasize that humans are novelty-seeking by nature (see, e.g., Clark, 2018; Gallagher, 2022). Here, we did not exploit this aspect of human nature by allowing discovery. It is the case that objects that were more surprising were considered more interesting than those that were not, suggesting that surprise still has an epistemic value; however, perhaps what it lacked was personal significance or the possibility to engage in further embodied exploration. Wohl (2022, p. 7) reports that "artists feel excitement when they are surprised by material results and feel that these results are relevant to their creative visions, they feel ambivalence when they are surprised by material results but are unclear as to the relevance to their creative visions." For surprise to have an effect, it is possible that it has to be meaningful. This echoes the findings from Ritter et al. (2012) that cognitive flexibility was only enhanced in situations where there was an active engagement with the schema violations through virtual reality. This would also lend support to theories that surprise has an effect because the human agent is motivated to find an explanation for the unexpected data point. Future

Table 5
Average (Standard Deviation) Originality as a Function of Believability and Congruence

Measure	Congruent	Incongruent
Believable	2.50 (0.76)	2.50 (0.78)
Unbelievable	2.53 (0.76)	2.58 (0.70)

[6] We would like to thank one of the reviewers of this article for this valuable observation.

investigations in the relationship between surprise and creativity should now focus on the role of a motivating affective response. This would link it more closely to existing findings associating creativity with unusual or impossible experiences.

The current paradigm also did not allow us to investigate whether the process of meaning-making (presumably sparked by the meaningfulness of the surprising observation) and schema updating is an important aspect in the relationship between surprise and creativity. When scientists report that they have made discoveries precipitated by a surprising and theoretically unexpected event, it is the updating of the theories that is considered to be the creative act. This is a broader implication than the simple disruption of perceptual schemas; originality in this sense would require not coming up with alternative uses for the object but plausible explanations for why the object was placed as it was. For example, Csikszentmihalyi (1996) tells the story of Vera Rubin's discovery of the galaxy rotation problem, which came when she was playing with overlapping pictures. The overlap was unexpected and surprising, but the recognition of the importance came from many years' attunement to typical patterns that allowed her to identify nontypical ones and also provided her with the skills to update the existing models. It is this metacognitive surprise that may drive creativity, particularly in the scientific domain. Such serendipitous findings are driven not only by the surprising observation but by the interpretation of that observation by a domain expert (Arfini & Costa, 2025; Ross, 2024; Yaqub, 2018).

In a related limitation, the use of a divergent thinking task limited the findings to the moment of initial idea generation. It is established that surprise is an important marker of what ideas are considered creative and may underlie the audience's assessment of creativity (Becattini et al., 2017). It may also be that the ideas that are generated by a surprising event may be more likely to elicit a feeling of surprise in the audience. Similarly, creativity consists of more than ideation. It may be that ideas which are generated by surprise also generate positive affect and metacognitive markers of confidence, which may make them something that would be more likely to be followed up. This would align with theories that suggest the "Aha" feeling functions as a metacognitive marker highlighting the value of an idea (Laukkonen et al., 2022). Future research could assess if the feeling of surprise prior to generating an idea leads participants to value the idea more than an idea with a similar level of originality.

Finally, the current set of experiments took place online. An increasing amount of psychological research is taking place online (Tomczak et al., 2023), but it is currently unclear whether data collected in this way should be considered as isomorphic to data collected in the laboratory. The benefits of a large and diverse participant pool who are more likely to be naïve to the hypotheses than psychology undergraduates may be offset by the requirements for focus in this task. If the effect of a surprising event requires immersion in the situation, then it is likely that the environment would need to be more controlled. It is worth noting that the cognitive cost of pursuing a surprising event may mean that in low-value situations like a psychological experiment, any effects may be minimal. This interpretation of the data, that human agents are surprise averse in this kind of artificial situations, supports the argument put forward by Ormerod (2023) that people tend to generate fewer possibilities than one might expect. It is unclear whether this cost is particularly salient for online studies or if it would be relevant in the laboratory as well. Similarly, it is possible that the online context cues a habituated response to the incongruous pictures because they are often encountered in this space, although again, it is unclear whether this habituation would also occur in an offline context as well. Future explorations of the effects that emerged in the present work should use a more ecological approach to test these questions. Even if we adopted highly controlled experimental conditions, we would suggest immersing participants in more realistic situations through the use, for instance, of virtual reality or by creating ad hoc environments with several schema disruptions in object disposition or use. Under these conditions, we could assume that surprise could be felt by participants as less artificial, and a higher arousal level could emerge in these situations. It is also likely that these situations would be more meaningful for participants.

Conclusion

Taken together, the current set of experiments establishes that simple schema discrepancy does not appear to be sufficient to trigger original uses for objects. This then establishes a methodological and theoretical baseline for future inquiry. These findings challenge assumptions about the relationship between surprise and creativity that permeate popular understanding. The data suggest that the folk wisdom linking surprise to creative inspiration may be overly simplistic, at least when surprise is conceived of as basic schema violations.

Given the existing theoretical assumptions about the relationship between surprise and creativity, we suggest that this is taken as a baseline for future research in this area. Future investigations should focus on identifying the specific conditions under which surprise does facilitate creativity, potentially emphasizing the roles of agency, meaningfulness, embodied exploration, and genuine affective engagement in transforming surprising encounters into creative insights.

References

Agnoli, S., Franchin, L., Rubaltelli, E., & Corazza, G. E. (2015). An eye-tracking analysis of irrelevance processing as moderator of openness and creative performance. *Creativity Research Journal*, *27*(2), 125–132. https://doi.org/10.1080/10400419.2015.1030304

Arfini, S., & Costa, M. (2025). What kind of explanations can serendipity provide? *European Journal for Philosophy of Science*, *15*(3), Article 42. https://doi.org/10.1007/s13194-025-00670-4

Baas, M., De Dreu, C. K. W., & Nijstad, B. A. (2008). A meta-analysis of 25 years of mood-creativity research: Hedonic tone, activation, or regulatory focus? *Psychological Bulletin*, *134*(6), 779–806. https://doi.org/10.1037/a0012815

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

Beaty, R., & Johnson, D. R. (2020). *Automating creativity assessment with SemDis: An open platform for computing semantic distance* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/nwvps

Becattini, N., Borgianni, Y., Cascini, G., & Rotini, F. (2017). Surprise and design creativity: Investigating the drivers of unexpectedness. *International Journal of Design Creativity and Innovation*, *5*(1–2), 29–47. https://doi.org/10.1080/21650349.2015.1090913

Benedek, M., Schües, T., Beaty, R. E., Jauk, E., Koschutnig, K., Fink, A., & Neubauer, A. C. (2018). To create or to recall original ideas: Brain processes associated with the imagination of novel object uses. *Cortex*, *99*, 93–102. https://doi.org/10.1016/j.cortex.2017.10.024

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. Routledge.

Brito, M. (2021, November 23). The infallible link between surprise and creativity. *The Groove*. https://blog.mariabrito.com/articles/the-groove-issue-62-the-infallible-link-between-surprise-and-creativity

Brunelle, T. (2021, June 27). Harnessing creativity and innovation: The value of surprise. *Medium*. https://tbrunelle.medium.com/harnessing-creativity-and-innovation-the-value-of-surprise-53637cb9b969

Chirico, A., Carrara, S., Bastoni, S., Gianotti, E., & Gaggioli, A. (2020). The effects of an ecological diversifying experience on creativity: An experimental study. *Frontiers in Psychology*, *11*(13), Article 1396. https://doi.org/10.3389/fpsyg.2020.01396

Chirico, A., Glaveanu, V. P., Cipresso, P., Riva, G., & Gaggioli, A. (2018). Awe enhances creative thinking: An experimental study. *Creativity Research Journal*, *30*(2), 123–131. https://doi.org/10.1080/10400419.2018.1446491

Chronicle, E. P., Ormerod, T. C., & MacGregor, J. N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *The Quarterly Journal of Experimental Psychology Section A*, *54*(3), 903–919. https://doi.org/10.1080/713755996

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, *17*(3), 521–534. https://doi.org/10.1007/s11097-017-9525-z

Corazza, G. E. (2023). Beyond the adjacent possible: On the irreducibility of human creativity to biology and physics. *Possibility Studies & Society*, *1*(1-2), 37–45. https://doi.org/10.1177/27538699221145664

Costa, P. T., & McCrae, R. R. (2008). *The revised Neo Personality Inventory*. Sage.

Csikszentmihalyi, M. (1996). *Creativity: The psychology of disovery and invention*. Harper Collins.

Currie, A. (2018). The argument from surprise. *Canadian Journal of Philosophy*, *48*(5), 639–661. https://doi.org/10.1080/00455091.2017.1368860

Damian, R. I. (2025). Diversifying experiences, personality, and creativity. *Creativity Research Journal*, *37*(2), 207–214. https://doi.org/10.1080/10400419.2024.2436318

Danek, A. H., Fraps, T., von Maller, A., Grothe, B., & Ollinger, M. (2014). It's a kind of magic: What self-reports can reveal about the phenomenology of insight problem solving. *Frontiers in Psychology*, *5*, Article 1408. https://doi.org/10.3389/fpsyg.2014.01408

Danek, A. H., Williams, J., & Wiley, J. (2020). Closing the gap: Connecting sudden representational change to the subjective Aha! Experience in insightful problem solving. *Psychological Research*, *84*(1), 111–119. https://doi.org/10.1007/s00426-018-0977-8

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Filipowicz, A. (2006). From positive affect to creativity: The surprising role of surprise. *Creativity Research Journal*, *18*(2), 141–152. https://doi.org/10.1207/s15326934crj1802_2

Foster, M. I., & Keane, M. T. (2019). The role of surprise in learning: Different surprising outcomes affect memorability differentially. *Topics in Cognitive Science*, *11*(1), 75–87. https://doi.org/10.1111/tops.12392

French, S., & Murphy, A. (2023). The value of surprise in science. *Erkenntnis*, *88*(4), 1447–1466. https://doi.org/10.1007/s10670-021-00410-z

Gallagher, S. (2022). Surprise! Why enactivism and predictive processing are parting ways: The case of improvisation. *Possibility Studies & Society*, *1*(3), 269–278. https://doi.org/10.1177/27538699221132691

George, T., Mielicki, M. K., & Wiley, J. (2023). Great expectations: Misleading effects of images in the alternate uses task. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(1), 56–67. https://doi.org/10.1037/aca0000380

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. https://doi.org/10.1111/j.2044-8295.2007.tb00467.x

Gonzalez, M. E. Q., & Haselager, W. F. G. (2005). Creativity: Surprise and abductive reasoning. *Semiotica*, *2005*(153-1/4), 325–341. https://doi.org/10.1515/semi.2005.2005.153-1-4.325

Green, A. E., Beaty, R. E., Kenett, Y. N., & Kaufman, J. C. (2023). The process definition of creativity. *Creativity Research Journal*, *36*(3), 544–572. https://doi.org/10.1080/10400419.2023.2254573

Guildford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. McGraw-Hill.

Hanchett Hanson, M. (2023). Reality studies: Reflections on the possible and actual. *Possibility Studies & Society*, *1*(1–2), 87–98. https://doi.org/10.1177/27538699231162311

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z., Sakhadeo, U. R., Baron, A. E., & Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software*, *54*(10), 1–26. https://doi.org/10.18637/jss.v054.i10

Laukkonen, R. E., Webb, M. E., Salvi, C., Tangen, J. M., Slagter, H. A., & Schooler, J. (2022). *Eureka heuristic: How feelings of insight signal the precision of a new idea* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/ez3tn

Lins de Holanda Coelho, G., Hanel, P. H. P., & Wolf, L. J. (2020). The very efficient assessment of need for cognition: Developing a six-item version. *Assessment*, *27*(8), 1870–1885. https://doi.org/10.1177/1073191118793208

Malthouse, E., Liang, Y., Russell, S., & Hills, T. (2022). The influence of exposure to randomness on lateral thinking in divergent, convergent, and creative search. *Cognition*, *218*, Article 104937. https://doi.org/10.1016/j.cognition.2021.104937

Mastria, S., Agnoli, S., Corazza, G. E., Grassi, M., & Franchin, L. (2022). What inspires us? An experimental analysis of the semantic meaning of irrelevant information in creative ideation. *Thinking & Reasoning*, *29*(4), 698–725. https://doi.org/10.1080/13546783.2022.2132289

Mealor, A. D., Simner, J., Rothen, N., Carmichael, D. A., Ward, J., & Schendan, H. (2016). Different dimensions of cognitive style in typical and atypical cognition: New evidence and a new measurement tool. *PLOS ONE*, *11*(5), Article e0155483. https://doi.org/10.1371/journal.pone.0155483

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220–232. https://doi.org/10.1037/h0048850

Meyer, W.-U., Niepel, M., Rudolph, U., & Schützwohl, A. (1991). An experimental analysis of surprise. *Cognition & Emotion*, *5*(4), 295–311. https://doi.org/10.1080/02699939108411042

Munnich, E. L., Foster, M. I., & Keane, M. T. (2019). An appraisal of surprise: Tracing the threads that stitch it together. *Topics in Cognitive Science*, *11*(1), 37–49. https://doi.org/10.1111/tops.12402

Neta, M., & Kim, M. J. (2023). Surprise as an emotion: A response to Ortony. *Perspectives on Psychological Science*, *18*(4), 854–862. https://doi.org/10.1177/17456916221132789

Noordewier, M. K., Topolinski, S., & Van Dijk, E. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass*, *10*(3), 136–149. https://doi.org/10.1111/spc3.12242

Öhlschläger, S., & Võ, M. L.-H. (2017). SCEGRAM: An image database for semantic and syntactic inconsistencies in scenes. *Behavior Research Methods*, *49*(5), 1780–1791. https://doi.org/10.3758/s13428-016-0820-3

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. Cambridge University Press.

Organisciak, P., Acar, S., Dumas, D., & Berthiaume, K. (2023). Beyond semantic distance: Automated scoring of divergent thinking greatly improves with large language models. *Thinking Skills and Creativity*, *49*, Article 101356. https://doi.org/10.1016/j.tsc.2023.101356

Ormerod, T. C. (2023). Possible, yes, but useful? Why the search for possibilities is limited but can be enhanced by expertise. *Possibility Studies & Society*, *1*(1-2), 178–189. https://doi.org/10.1177/27538699231172562

Ortony, A. (2022). Are all "basic emotions" emotions? A problem for the (basic) emotions construct. *Perspectives on Psychological Science*, *17*(1), 41–61. https://doi.org/10.1177/1745691620985415

Ortony, A., & Russell, J. A. (2024). On the invalidity of Neta and Kim's argument that surprise is always valenced. *Emotion Review*, *16*(1), 64–67. https://doi.org/10.1177/17540739231214785

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perkins, D. N. (2001). *The eureka effect: The art and logic of breakthrough thinking*. Norton.

Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 144–152. https://doi.org/10.1037/aca0000227

Richardson, D. C., Tagomori, H., & Devlin, J. T. (2022). Can extreme experiences enhance creativity? The case of the underwater nightclub. *Frontiers in Psychology*, *13*, Article 785278. https://doi.org/10.3389/fpsyg.2022.785278

Ritter, S. M., Damian, R. I., Simonton, D. K., van Baaren, R. B., Strick, M., Derks, J., & Dijksterhuis, A. (2012). Diversifying experiences enhance cognitive flexibility. *Journal of Experimental Social Psychology*, *48*(4), 961–964. https://doi.org/10.1016/j.jesp.2012.02.009

Ross, W. (2024). Accidental thinking: A model of serendipity's cognitive processes. *Review of General Psychology*, *28*(3), 253–267. https://doi.org/10.1177/10892680241254759

Ross, W., & Arfini, S. (2023). Serendipity and creative cognition. In L. J. Ball & F. Vallée-Tourangeau (Eds.), *Routledge handbook of creative cognition* (pp. 35–54). Routledge.

Ross, W., & Vallée-Tourangeau, F. (2022). Accident and agency: A mixed methods study contrasting luck and interactivity in problem solving. *Thinking & Reasoning*, *28*(4), 487–528. https://doi.org/10.1080/13546783.2021.1965025

Shir, Y., Abudarham, N., & Mudrik, L. (2021). You won't believe what this guy is doing with the potato: The ObjAct stimulus-set depicting human actions on congruent and incongruent objects. *Behavior Research Methods*, *53*(5), 1895–1909. https://doi.org/10.3758/s13428-021-01540-6

Sundquist, D., & Lubart, T. (2022). Being intelligent with emotions to benefit creativity: Emotion across the seven cs of creativity. *Journal of Intelligence*, *10*(4), Article 106. https://doi.org/10.3390/jintelligence10040106

Tomczak, J., Gordon, A., Adams, J., Pickering, J. S., Hodges, N., & Evershed, J. K. (2023). What over 1,000,000 participants tell us about online research protocols. *Frontiers in Human Neuroscience*, *17*, Article 1228365. https://doi.org/10.3389/fnhum.2023.1228365

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*(1), 1–40. https://doi.org/10.1006/cogp.1994.1010

Williams, D. (2015). *Leadership for a fractured world: How to cross boundaries, build bridges, and lead change* (1st ed.). Berrett-Koehler Publishers.

Wiseman, R., & Watt, C. (2022). Experiencing the impossible and creativity: A targeted literature review. *PeerJ*, *10*, Article e13755. https://doi.org/10.7717/peerj.13755

Wohl, H. (2022). Innovation and creativity in creative industries. *Sociology Compass*, *16*(2), Article e12956. https://doi.org/10.1111/soc4.12956

Yaqub, O. (2018). Serendipity: Towards a taxonomy and a theory. *Research Policy*, *47*(1), 169–179. https://doi.org/10.1016/j.respol.2017.10.007

Received January 10, 2025
Revision received September 24, 2025
Accepted October 5, 2025 ■