Personality and Individual Differences 189 (2022) 111473

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Examples in creative exhaustion: The role of example features and individual differences in creativity

Huan Yuan [a,1], Kelong Lu [b,1], Mengsi Jing [c], Cuirong Yang [a,*], Ning Hao [b,**]

[a] School of Education, Suzhou University of Science and Technology, Suzhou, China
[b] Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China
[c] School of Education, Suzhou University, Suzhou, China

ARTICLE INFO

Keywords:
Exhaustion
Example
Novelty
Diversity
Individual differences

ABSTRACT

This study examined how example features (i.e. idea novelty and idea diversity) affected creativity after creative exhaustion. In addition, how individual differences in creativity influenced the effect of example diversity was also investigated. By dividing examples into high-, medium- and low-novel ideas, experiment 1 revealed that idea originality was more enhanced in the high-novel example group than in the low-novel example and none example groups (control group). The low-novel group showed higher idea fluency and flexibility than the control group. In experiment 2, participants were divided into high-creative and low-creative groups. Both groups received diverse and similar examples with high novelty. Results showed that idea fluency, originality and flexibility were higher in the high-creative group than those in the low-creative group. Moreover, diverse examples facilitated fluency and flexibility while similar examples facilitated originality after creative exhaustion. These findings indicate that example features affect the stimulation effect of example after creative exhaustion.

1. Introduction

Creativity is defined as the ability to generate original and useful ideas, insights, or problem solutions (Amabile, 1983; Sternberg & Lubart, 1999). People often encounter mental fixation in creativity practice, which is defined as an inappropriate adherence to an approach in problem solving or generating creative ideas for more open-ended tasks (Koppel & Storm, 2014; Smith, 1995). During creative idea generation, a similar scenario is creative exhaustion, which depicts an inability to continue generating creative solutions on one's own (Gray et al., 2019). Creative exhaustion typically occurs when people exhaust their initial ideas. It impedes creative idea generation process and leads to a gap where new creative ideas can hardly be generated. Therefore, it is of high significance to deal with creative exhaustion and enable creative idea generation to continue running.

Previous studies have demonstrated that cognitive stimulation such as providing cognitive supports could facilitate idea generation after creative exhaustion. For example, Gray et al. (2019) reported that design tools could inspire beginning designers to generate more ideas after creative exhaustion. As another sort of cognitive stimulation, exposing individuals to existing ideas (or examples) could also successfully stimulate creativity regardless of whether encountering exhaustion or not (Althuizen & Wierenga, 2014; Fink et al., 2012; Lambert et al., 2019; Nijstad & Stroebe, 2006; Yuan et al., 2021). For instance, presenting existing example solutions can inspire more novel ideas after idea exhaustion (Siangliulue et al., 2015). However, previous creative exhaustion studies only presented examples with high novelty and diversity to participants. How example features (e.g. different levels of novelty and diversity) can affect the stimulation effect of example (example effect) after creative exhaustion is still unknown.

Idea novelty or originality is one of the key features affecting the example effect (Fink et al., 2012; George & Wiley, 2020; Pi et al., 2019). According to the Search for Ideas in Associative Memory (SIAM) model, idea generation is a repeated search process in associative memory (Nijstad & Stroebe, 2006). Examples serving as the search cues will activate problem-relevant knowledge, which will further facilitate idea

* Correspondence to: C. Yang, School of Education, Suzhou University of Science and Technology, No.1, Kerui Road, Suzhou 215009, China.
** Correspondence to: N. Hao, Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, No. 3663, North Zhong Shan Road, Shanghai 200062, China.
E-mail addresses: yangsun2004@mail.usts.edu.cn (C. Yang), nhao@psy.ecnu.edu.cn (N. Hao).
[1] These authors contributed equally.

https://doi.org/10.1016/j.paid.2021.111473
Received 22 October 2021; Received in revised form 9 December 2021; Accepted 14 December 2021
Available online 5 January 2022

generation. Note that examples can be either novel or common, which may affect the categories of problem-relevant knowledge that are activated. This may eventually moderate the example effect on creative idea generation. For example, George et al. (2019) presented three toy examples to participants during a toy designing task and found that moderately novel examples facilitated the novelty of generated ideas. According to the classic associative model of creativity, the activation of remote concepts needs more time in associative memory (Mednick, 1962). Therefore, the creative exhaustion may be along with activating remote concept in associative network after exhausting initial ideas. As the serial order effect suggests, common ideas usually appear in the early stage of creative ideation (Christensen et al., 1957). Gilhooly et al. (2007) also found that initial ideas were retrieved from long term memory of existing knowledge and were of low novelty. Accordingly, common examples may help activate close associations that have been already utilized before creative exhaustion. In case of creative exhaustion, being exposed to novel examples rather than common ones may help activate remote associations, which have not been activated due to their remoteness to the task stimulus, and breed more novel ideas.

Likewise, idea diversity was also confirmed to influence the example effect. Idea diversity refers to the semantic distance between ideas or number of categories the ideas belong to (Jonathan et al., 2021). Previous studies have shown that early exposure to diverse ideas increases the novelty of the final product (Baruah & Paulus, 2011; Jonathan et al., 2021). According to the SIAM model, diverse stimuli serving as search cues can activate different problem-relevant knowledge from one's memory. For example, Nijstad et al. (2002) have found diverse stimuli could facilitate participants generating a greater variety of solutions. In other words, diverse stimuli contribute to idea fluency and flexibility. According to the open goal effect, when individual temporally breaks from an unsolved problem, being exposed to an implicit hint can even help solve the problem. That is, open goals influence the acquisition of problem-relevant information. Creative exhaustion is the moment that people realize that they have exhausted ideas. In case of creative exhaustion, it is possible that individuals will be ready to use external stimuli to generate new ideas after running out of ideas (Moss et al., 2007). However, according to the attention allocation theory, presenting examples can direct individual attention to example-related domains and thus initiate deeper explorations in these domains. As diverse examples will activate diverse relevant knowledge domain, and direct individuals to explore potential ideas across these domains. Such a more diffused search strategy may attenuate individual tendency to deeply explore ideas within each domain and thereby reduce idea originality (Sio et al., 2015).

Moreover, individual differences in creativity may also moderate the example effect. According to Mednick's (1962) theory, creative individuals have a richer and more flexible associative network than less creative ones. Kenett et al. (2014) extended Mednick's theory by using network analysis approach to explore semantic memory networks and further suggested that the semantic memory network of less creative individuals seems to be more rigid than that of creative individuals. Moreover, ideas are commonly considered as products of associative memories (Lambert et al., 2019). We suggested that examples can activate more associations in a richer and more flexible associative network than in a more rigid one. Therefore, creative individuals may be more flexible and efficient to search example-related ideas (especially for diverse ideas) from their associative networks and thereby perform better than less creative ones after creative exhaustion.

To date on, how example features (i.e. novelty and diversity) and individual differences in creativity affect the example effect after creative exhaustion is less investigated. This study aimed to address the following questions: (1) How does idea (example) novelty and diversity affect the example effect after creative exhaustion? (2) How do individual differences in creativity affect the example diversity effect after creative exhaustion?

The current study tested how example novelty (Experiment 1) and example diversity (Experiment 2) affected creative idea generation after creative exhaustion. In addition, whether high- and low-creative people showed difference on the effect of example diversity was also investigated in experiment 2. Previous studies have only dichotomized existing ideas into common and novel ones based on their novelty (Fink et al., 2012; Pi et al., 2019; Wang et al., 2018). By further dividing these example ideas into high-, medium- and low-novel (common) ideas, the current study aimed to investigate which kind of examples could boost creativity most after creativity exhaustion. As one of the key components of creativity, divergent thinking (DT) is predictive to creative potential (Runco & Acar, 2012). Therefore, the current study utilized DT task (i.e. alternative uses task, AUT) to measure creativity. Creative performance was evaluated by independent ratings of fluency, originality and flexibility (Guilford, 1967; Runco & Jaeger, 2012; Runco & Pritzker, 1999). Based on the SIAM theory, associative theory of creativity and attention allocation explanation aforementioned, we proposed three hypotheses: H1: High-novel examples will boost fluency, originality and flexibility most after creative exhaustion; H2: Diverse examples will stimulate higher fluency and flexibility than similar examples, but similar examples will facilitate higher originality than diverse examples after creative exhaustion; H3: When being exposed to examples (especially diverse ones) after creative exhaustion, the high-creative group will perform better than the low-creative group.

2. Experiment 1

Experiment 1 investigated the effect of example novelty on DT after creative exhaustion. Previous studies have shown common and novel examples could function differently during idea generation (Fink et al., 2012; Perttula & Sipilä, 2007; Pi et al., 2019; Sio et al., 2015). However, the role of example novelty has not yet been explored in creative exhaustion. Thus, participants were asked to generate as many and original solutions as possible without time limitation so as to experience creative exhaustion. Examples varying in novelty (i.e. high novelty, medium novelty and low novelty) were presented to participants after creative exhaustion. Participants received no example in the control group.

2.1. Participants and design

A priori power analysis using G*power 3.1 (Faul et al., 2007) was conducted to estimate the sample size necessary for the main effect at 80% power. The between participants effect size was set to $\eta^2 = 0.25$. According to the results, the required sample size is 96. We recruited 97 senior high school students. A one-factor (Example Novelty: high-novel examples, medium-novel examples, low-novel examples vs. control) between-subjects design was employed. Participants were randomly assigned to one of the four groups.

Among the 97 participants, one was excluded because he failed to complete the experiment as requested. Thus, the final sample consisted of 96 participants (51 females, 44 males; age: 16.97 $\pm$ 0.31 years old). There were respectively 25, 25, 24 and 22 participants in the low-novel, medium-novel, high-novel and control groups. All participants were right-handed and native speakers of Chinese. They signed the written informed consent prior to the experiment and received approximately 4 US dollars for their participation. The experiment protocol was approved by the Institutional Ethics Committee of the University.

2.2. Task and procedure

This study was conducted during the outbreak period of COVID-19 when senior high school students were asked to learn online at home in China. Therefore, the study was carried out online through a messaging app WeChat. To ensure task commitment and valid performance, participants were asked to find a quiet room without disturbance before the experiment, and turn on the webcam so that the whole

experiment proceeded under supervision. Participants in each group were required to complete two AUT items (i.e. brick and newspaper). The presenting sequence of the AUT items were counterbalanced across participants. AUT is a well-established test of creative potential (Guilford, 1967; Runco & Mraz, 1992). It requires participants to generate as many and original uses as possible for common objects. The introduction of AUT was "*Please produce as many and original uses as you can think of, which are different from the normal use, for two common objects without time limitation. When you run out of ideas, press 'can't' into the computer*". Each task was divided into two parts: a free ideation stage to induce creative exhaustion (stage 1) and a continued creative ideation stage (stage 2). In stage 1, participants were asked to generate as many and original ideas as possible and type them into the computer without time limitation. Stage 2 began when participants reported that they had exhausted ideas. Two examples were then presented to participants in the form of text in the example groups. After understanding the examples, they were asked to continue thinking for 5 min. Participants were not instructed that they would get to work on after exhaustion so that they could spare no effort to exhaust their ideas during the pre-exhaustion session. Participants in the control group continued thinking without receiving any hint (see Fig. 1). Except for whether being exposed to examples, the experimental manipulation did not differ between the example groups and control group. In order to ensure participants' attention on the examples, participants in the example groups were told that they needed to recall these examples after the experiment (Nijstad et al., 2002).

The examples used in experiment 1 were obtained from the original data of one prior research (Hao et al., 2014). According to the consensus assessment technique (CAT) on a 5-point Likert scale, ideas were divided into low- (originality scores ≤ 2), medium- ($2.5 \leq$ originality scores ≤ 3.5) and high-novel examples (originality scores ≥ 4).

2.3. Task assessment

Performance of AUT was evaluated in terms of fluency, originality and flexibility (Guilford, 1967; Runco & Pritzker, 1999). Fluency was measured as the number of non-redundant ideas that were produced by the participants. For originality, six raters independently scored each idea using a 5-point Likert scale (1 = unoriginal, 5 = original). The inter-rater reliability of this method was satisfactory (newspaper: $\alpha = 0.967$; brick: $\alpha = 0.928$). The final originality scores for each participant were obtained by averaging the individual ratings from the raters. We further assessed the consistency of originality scores of the two AUT items, and found response originality of brick correlated well with that of newspaper in the pre-exhaustion [r (96) = 0.36, $p < 0.001$], and post-exhaustion sessions [r (96) = 0.25, $p = 0.026$]. For flexibility, responses for each AUT item were grouped into 7 broad semantic categories. For instance, the response categories of *newspaper* included reading, handwork like paper folding, sheltering from things, decoration, stationery, cleaning tools, and others. The response categories of *brick* included constructing buildings, a tool of pounding, furniture like tables, physical attack or self-defense, decoration, playing games and others. Responses pertaining to same semantic categories were grouped. Two trained raters independently coded a random subset (25%) of the response pool. The inter-rater reliability (Cronbach's α) on these scores was satisfactory (newspaper: $\alpha = 0.792$; brick: $\alpha = 0.883$). Next, the first rater scored the remaining responses. The final flexibility scores for each participant were calculated by counting the number of explored categories. Note that raters of originality and flexibility were blind to experimental conditions.

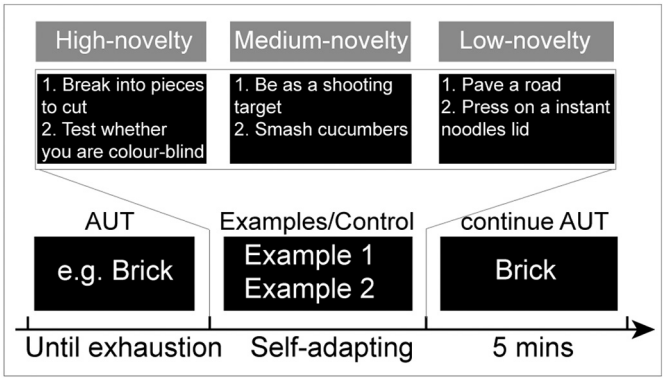

Fig. 1. Procedure overview of one of the two AUT items. *Note*: Participants worked on the AUT until they reported that they had run out of ideas (i.e. creative exhaustion). Those in the Example groups were then exposed to two examples with high/medium/low novelty. The control group received no example and continued AUT for 5 min. The example groups would continue AUT for 5 min after understanding the examples by self-adaption.

2.4. Post-experiment tests

After the experiment, participants were asked to recall the examples presented during the tasks. Moreover, they were asked to rate the examples' usefulness and their usage value on a 5-likert scale (1 = not at all, 5 = very much). For example, they were asked to complete the items such as: "*How useful are these examples for you?*" and "*How much do you consult these examples during the task?*"

2.5. Results

A one-way ANOVA or MANOVA with Example Novelty (high-novel vs. medium-novel vs. low-novel vs. control group) as the between-subject factor was performed on each dependent variable (i.e. task duration, pre-exhaustion AUT performance, and post-exhaustion AUT performance). All post hoc tests were corrected using the Scheffe correction in this study.

2.5.1. Pre-exhaustion task duration in different groups

Shapiro–Wilk (S-W) tests were used to examine the normality of the task duration and showed that task duration in the medium-novel ($p = 0.001$) and control groups ($p = 0.041$) did not normally distribute. Therefore, a Kruskal-Wallis (K-W) test was performed to compare the pre-exhaustion task duration between groups and revealed no significant difference ($p = 0.075$).

2.5.2. Pre-exhaustion AUT performance in different groups

MANOVAs on pre-exhaustion AUT performance (fluency, originality and flexibility) (*Box*'s $M = 21.487$, $p = 0.321$) revealed a significant main effect of Example Novelty on originality, F (3, 92) = 7.99, $p < 0.001$, $\eta_p^2 = 0.21$. Post hoc tests showed originality was higher in the high-novel group ($M = 1.84$, $SD = 0.33$; $p = 0.001$, *Cohen*'s $d = 1.36$) and medium-novel group ($M = 1.86$, $SD = 0.37$; $p < 0.001$, *Cohen*'s $d = 1.31$) than in the control group ($M = 1.47$, $SD = 0.20$). No significant main effect of Example Novelty was observed on fluency [F (3, 92) = 0.122, $p = 0.947$] and flexibility [F (3, 92) = 0.119, $p = 0.948$] (see Fig. 2).

2.5.3. The effect of example novelty on post-exhaustion AUT performance

One-way MANOVA showed that the covariance matrices of post-exhaustion creative performance indices (fluency, originality and flexibility) were not homogeneous (*Box*'s $M = 21.487$, $p = 0.005$), thereby the data did not fit MANOVA. Thus, one-way ANOVAs using Example Novelty as the between-group factor were conducted on the above dependent variables, respectively.

Results showed a significant main effect of Example Novelty on fluency, F (3, 92) = 3.36, $p = 0.022$, $\eta_p^2 = 0.10$. Post hoc tests showed that fluency in the low-novel group ($M = 4.08$, $SD = 1.78$) was higher than that in the control group ($M = 2.36$, $SD = 1.54$, $p = 0.028$, Cohen's $d = 1.03$). The main effect of Example Novelty on originality was significant, F (3, 92) = 17.29, $p < 0.001$, $\eta_p^2 = 0.36$. Post hoc tests revealed

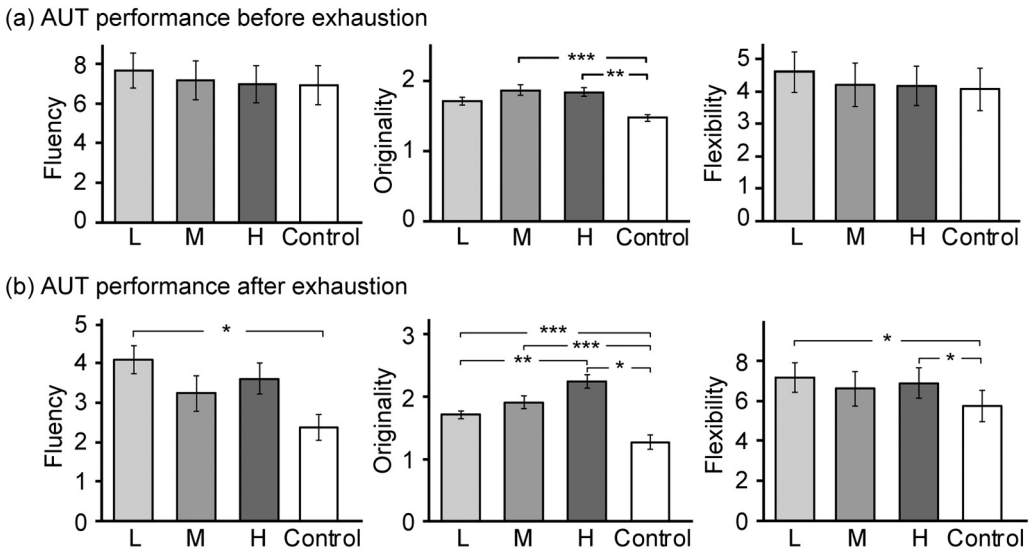

Fig. 2. (a) and (b) refer to the fluency, originality and flexibility of responses in the high-novel (H), medium-novel (M), low-novel (L) groups and control group before and after creative exhaustion separately. Error bars indicate standard errors of the mean. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

that originality in the high-novel group ($M = 2.24$, $SD = 0.54$, $p = 0.019$, Cohen's $d = 1.83$), medium-novel group ($M = 1.91$, $SD = 0.48$, $p < 0.001$, Cohen's $d = 1.23$), and low-novel group ($M = 1.71$, $SD = 0.30$, $p < 0.001$, Cohen's $d = 1.04$) were higher than that in the control group ($M = 1.27$, $SD = 0.52$). Moreover, originality in the high-novel group ($M = 2.24$, $SD = 0.54$) was higher than that in the low-novel group ($M = 1.71$, $SD = 0.30$, $p = 0.002$, Cohen's $d = 1.21$). The main effect of Example Novelty on flexibility was significant, F (3, 92) = 4.44, $p = 0.006$, $\eta_p^2 = 0.13$. Post hoc tests showed that flexibility in the high-novel group ($M = 2.58$, $SD = 1.14$, $p = 0.015$, Cohen's d = 1.03) and low-novel group ($M = 2.48$, $SD = 0.92$, $p = 0.033$, Cohen's $d = 1.06$) were higher than that in the control group ($M = 1.55$, $SD = 0.83$) (see Fig. 2).

As pre-exhaustion originality differed among four groups, they were treated as covariates in the aforementioned ANOVA models. Main effects of Example Novelty on fluency, originality and flexibility remained significant [fluency: F (3, 91) = 2.74, $p = 0.045$, $\eta_p^2 = 0.08$; originality: F (3, 91) = 11.17, $p < 0.001$, $\eta_p^2 = 0.27$; flexibility: F (3, 91) = 3.24, $p = 0.026$, ηp2 = 0.10].

2.5.4. Example usefulness, usage, and example recall in different groups

A one-way MANOVA with Example Novelty as the between-subject factor was performed on example usefulness and example usage (*Box*'s $M = 13.12$, $p = 0.05$). A significant main effect of Example Novelty was observed on example usefulness, F (2, 71) = 4.49, $p = 0.015$, $\eta_p^2 = 0.11$. Post hoc tests revealed that example usefulness in the low-novel group ($M = 2.97$, $SD = 0.57$) was lower than that in the medium-novel group ($M = 3.28$, $SD = 0.96$, $p = 0.03$, Cohen's $d = -0.36$) and high-novel group ($M = 3.41$, $SD = 0.81$, $p = 0.05$, Cohen's $d = -0.63$). No significant main effect of Example Novelty was observed on example usage, F (2, 71) = 2.01, $p = 0.14$. Pearson correlation analysis showed that example usefulness was positively correlated with fluency ($r = 0.44$, $p < 0.001$), originality ($r = 0.29$, $p = 0.01$) and flexibility ($r = 0.48$, $p < 0.001$) after creative exhaustion. Moreover, all participants successfully recalled the examples.

2.6. Interim discussion

These results partly supported Hypothesis 1. High-novel examples were more beneficial to post-exhaustion DT performance than low-novel examples and no example. That is, the example effect on post-exhaustion DT performance depends on example novelty. Additionally, novel examples become more accessible over time.

High-novel examples facilitated creative ideation after creative exhaustion. This could be explained by the association between examples and the activated semantic network. As the SIAM theory argues, the proper cues can activate the associative memory and boost idea generation as a result (Nijstad & Stroebe, 2006). According to the serial order effect, individuals would generate more novel but less ideas in the late stage (Beaty & Silvia, 2012; Hass & Beaty, 2018; Wang et al., 2017). Considering these theories, remote areas of the semantic network are activated in the creative exhaustion stage. It is possible that high-novel examples were more likely to activate remote areas of the semantic network, which have not been activated due to their remoteness to the task stimulus, than low-novel ones and breed more original ideas.

3. Experiment 2

Based on the results of experiment 1 and SIAM theory, experiment 2 aimed to further investigate the effect of example diversity on post-exhaustion DT performance. In experiment 1, the examples were diverse rather than similar (share many characteristics among the stimuli themselves), and thus effect of example diversity on post-exhaustion DT performance remained unexplored. Moreover, findings in experiment 1 showed that high-novel examples facilitated idea originality most. Therefore, high-novel examples were used in experiment 2. In addition, how individual differences in creativity affect this example effect was also investigated.

3.1. Participants and design

The results of a priori power analysis using G*power 3.1 (Faul et al., 2007) showed the required sample size is 86. Before the experiment, one hundred and eighty-eight senior high school students completed the Creativity Tendency Scale (CTS) of Creativity Assessment Packet (CPA) and product improvement task (PIT). PIT is the subtest 4 of TTCT, where participants are required to think of as many ways as possible to change a toy to make it more enjoyable and appealing (Chen et al., 2016). According to the total scores of these two tasks, the top 23% and bottom 23% individuals were identified as high- and low-creative individuals, respectively.

Eighty-four senior high school students participated in experiment 2 (46 females, 38 males; age: 17.01 ± 0.4 years old). A 2 (Creativity Level: high vs. low) × 2 (Example Diversity: similar vs. diverse) mixed-subject design was employed with Creativity Level as the between-subject factor

and Example Diversity as the within-subject factor. The final sample sizes of the high-creative group and low-creative group were both 42.

All participants were right-handed and native speakers of Chinese. They signed the written informed consent prior to the experiment and received approximately 4 US dollars for their participation. The experiment protocol was approved by the Institutional Ethics Committee of the University.

3.2. Procedure

Similar to experiment 1, the formal experiment consisted of two stages. Stage 1 was same as that in experiment 1. In stage 2, participants were asked to continue completing two AUT items while two diverse or similar ideas were presented in the form of text during each AUT task. During the task, participants were asked to generate as many and novel ideas as possible in 5 min and type them into the computer. AUT items were also presented in a balanced order.

Examples (originality > 4 within a 5-point Likert scale) from diverse/similar categories (rated by 2 raters, $\alpha = 0.96$) in the prior research (Hao et al., 2014) were assigned to the diverse group or similar group. As idea diversity refers to the number of categories ideas belong to or the semantic distance between ideas (Jonathan et al., 2021), example diversity of these two groups was further rated on a 5-point Likert scale (1 = close semantic distance, 5 = far semantic distance) by 6 raters to ensure the diversity manipulation (Jonathan et al., 2021). Inter-rater reliability on these scores was satisfactory ($\alpha = 0.84$). Examples for the similar group and diverse group were selected based on the averaging semantic distance ratings from the six raters (newspaper: similar group: 0.7, diverse group: 2.2; brick: similar group: 1.2, diverse group: 2.3).

Specifically, the similar examples for newspaper and brick included "*Splicing the words in the newspaper as an anonymous letter or using it as a codebook to exchange information*" and "*drilling holes in the brick to raise bees/earth worms*", respectively. The diverse examples for newspaper and brick included "*making a table from compacted newspapers or using newspapers as stage props like snowflakes*" and "*hollowing out the brick to make a tissue box or playing mahjong with bricks*", respectively.

3.3. Task assessment

Same with experiment 1, AUT performance was evaluated in terms of fluency, originality and flexibility (Guilford, 1967; Runco & Pritzker, 1999). For originality, six raters scored each idea by using a 5-point Likert scale (1 = unoriginal, 5 = original). Inter-rater reliability on these scores was satisfactory (newspaper: $\alpha = 0.94$; brick: $\alpha = 0.92$). AUT flexibility was scored as in study 1 (newspaper: $\alpha = 0.700$; brick: $\alpha = 0.915$). Note that raters of originality and flexibility were blind to experimental conditions. These final originality scores for each participant were obtained by averaging the individual ratings from the raters.

3.4. Post-experimental tests

Immediately after the task, participants were required to recall the examples. They were also asked to evaluate the examples' usefulness and usage using a 5-point Likert scale (four questions). Example questions are listed as follows: "*How useful are these examples for you*" and "*How much do you consult these examples during the task?*"

3.5. Results

A mixed design ANOVA, with Example Diversity (diverse vs. similar) as the within-subject factor and Creativity Level (high vs. low) as the between-subject factor, was performed on each dependent measure (i.e. task duration, pre-exhaustion AUT performance, post-exhaustion AUT performance).

3.5.1. Pre-exhaustion task duration in different groups

The ANOVA showed a significant main effect of Creativity Level on the pre-exhaustion task duration, $F(1, 82) = 5.853$, $p = 0.018$, $\eta_p^2 = 0.067$. In particular, the high-creative group showed longer pre-exhaustion task duration ($M = 475.52$ s, $SD = 272.094$ s) than the low-creative group ($M = 351.134$ s, $SD = 189.75$ s). Main effect of Example Diversity and interaction effect were not significant, $F(1, 82) = 0.648$, $p = 0.423$; $F(1, 82) = 0.277$, $p = 0.600$.

3.5.2. Pre-exhaustion AUT performance in different groups

The ANOVA demonstrated a main effect of Creativity Level on fluency, $F(1, 82) = 27.317$, $p < 0.001$, $\eta_p^2 = 0.250$. Results showed higher fluency in the high-creative group ($M = 11.34$, $SD = 7.56$) than in the low-creative group ($M = 6.17$, $SD = 3.67$, $p < 0.001$, Cohen's $d = 0.87$). No effect of Example Diversity and interaction effect was observed on fluency, $F(1, 82) = 0.01$, $p = 0.92$; $F(1, 82) = 0.00$, $p = 0.99$. For originality, results showed no significant main effect of Creativity Level [$F(1, 82) = 0.84$, $p = 0.36$], Example Diversity [$F(1, 82) = 1.14$, $p = 0.29$] or interaction effect [$F(1, 82) = 0.90$, $p = 0.35$]. For flexibility, results revealed a significant main effect of Creativity Level, $F(1, 82) = 26.07$, $p < 0.001$, $\eta_p^2 = 0.24$. In particular, the high-creative group performed higher flexibility ($M = 7.465$, $SD = 5.759$) than the low-creative group ($M = 3.65$, $SD = 2.445$, Cohen's $d = 0.86$). No significant main effect of Example Diversity and interaction effect was found, $F(1, 82) = 0.001$, $p = 0.98$; $F(1, 82) = 0.03$, $p = 0.86$ (see Fig. 3).

To exclude the potential contaminative effect of pre-exhaustion task duration, it was taken as a covariate in the aforementioned models. For fluency and flexibility, the main effect of Creativity Level remained significant, $F(1, 80) = 2.288$, $p < 0.001$, $\eta_p^2 = 0.20$; $F(1, 80) = 19.225$, $p < 0.001$, $\eta_p^2 = 0.19$. However, main effect of Example Diversity on originality became significant, $F(1, 80) = 2.288$, $p = 0.027$, $\eta_p^2 = 0.06$. In particular, the diverse example group showed higher originality ($M = 4.08$, $SD = 2.36$) than the similar example group ($M = 3.31$, $SD = 1.90$, Cohen's $d = 0.36$).

3.5.3. Post-exhaustion AUT performance in different groups

Results showed significant main effect of Example Diversity and Creativity Level on fluency, $F(1, 82) = 5.45$, $p = 0.022$, $\eta_p^2 = 0.06$; $F(1, 82) = 5.68$, $p = 0.019$, $\eta_p^2 = 0.07$. Fluency in the diverse example condition ($M = 4.08$, $SD = 2.40$) was higher than that in the similar example condition ($M = 3.32$, $SD = 1.93$, Cohen's $d = 0.35$). Fluency in the high-creative group ($M = 4.10$, $SD = 2.46$) was higher than that in the low-creative group ($M = 3.30$, $SD = 1.73$, Cohen's $d = 0.38$). No significant interaction effect of Creativity Level and Example Diversity was observed, $F(1, 82) = 0.24$, $p = 0.63$ (see Fig. 3).

For originality, results showed significant main effect of Example Diversity and Creativity Level, $F(1, 82) = 4.68$, $p = 0.034$, $\eta_p^2 = 0.05$; $F(1, 82) = 6.93$, $p = 0.01$, $\eta_p^2 = 0.08$. Originality in the similar example condition ($M = 2.38$, $SD = 0.77$) was higher than that in the diverse example condition ($M = 2.13$, $SD = 0.74$, Cohen's $d = 0.33$). Originality in the high-creative group ($M = 2.40$, $SD = 0.79$) was higher than that in the low-creative group ($M = 2.11$, $SD = 0.70$, Cohen's $d = 0.39$). No significant interaction effect of Creativity Level and Example Diversity was observed, $F(1, 82) = 0.02$, $p = 0.89$ (see Fig. 3).

For flexibility, results showed significant main effect of Example Diversity and Creativity Level, $F(1, 82) = 8.28$, $p = 0.005$, $\eta_p^2 = 0.09$; $F(1, 82) = 4.49$, $p = 0.037$, $\eta_p^2 = 0.05$. Flexibility in the diverse example condition ($M = 2.89$, $SD = 1.54$) was higher than that in the similar example condition ($M = 2.30$, $SD = 1.17$, Cohen's $d = 0.43$). Flexibility in the high-creative group ($M = 2.82$, $SD = 1.58$) was higher than that in the low-creative group ($M = 2.48$, $SD = 1.03$, Cohen's $d = 0.25$). No significant interaction effect of Creativity Level and Example Diversity was observed, $F(1, 82) = 0.002$, $p = 0.96$ (see Fig. 3).

To exclude the potential contaminative effect of pre-exhaustion task duration, it was taken as a covariate in the aforementioned models. For fluency, originality and flexibility, the main effect of Creativity Level

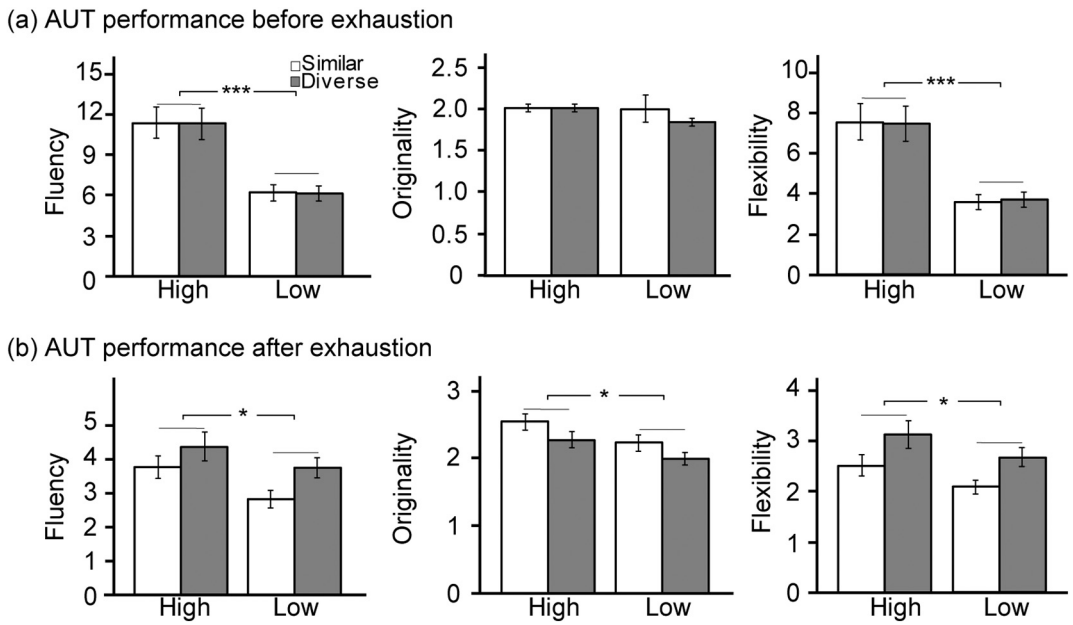

Fig. 3. (a) and (b) refers to refers to the fluency, originality and flexibility of responses before and after exhaustion, respectively. Error bars indicate standard errors of the mean. *$p < 0.05$, ***$p < 0.001$.

remained significant [fluency: F (1, 80) = 5.74, $p = 0.019$, $\eta_p^2 = 0.07$; originality: F (1, 80) = 4.32, $p = 0.04$, $\eta_p^2 = 0.05$; flexibility: F (1, 80) = 4.20, $p = 0.044$, $\eta_p^2 = 0.05$]. However, main effect of Example Diversity became non-significant, [fluency: F (1, 80) = 0.59, $p = 0.445$, originality: F (1, 80) = 2.65, $p = 0.11$, flexibility: F (1, 80) = 2.62, $p = 0.11$].

3.5.4. Example usefulness, usage, and example recall in different groups

There was no significant main effect of Creativity Level and Example Diversity and interaction effect between them on evaluations of example usefulness [Creativity Level: F (1, 81) = 1.36, $p = 0.25$; Example Diversity: F (1, 81) = 1.36, $p = 0.25$; Creativity Level × Example Diversity: F (1, 81) = 0.18, $p = 0.67$]. Similarly, results showed no significant main effect of Creativity Level and Example Diversity and interaction effect between them on example usage [Creativity Level: F (1, 81) = 1.36, $p = 0.25$; Example Diversity: F (1, 81) = 0.17, $p = 0.68$; Creativity Level × Example Diversity: F (1, 81) = 0.51, $p = 0.48$]. Moreover, all participants successfully recalled the examples.

3.6. Interim discussion

Results showed that post-exhaustion fluency and flexibility were higher in the diverse example condition than those in the similar example condition. In contrast, originality benefited more from similar examples than diverse examples. This result supported Hypothesis 2. When being exposed to examples (regardless of example diversity), high-creative participants performed well in terms of idea fluency, originality, and flexibility when compared to those with low creativity. Hypothesis 3 was partly supported. However, diverse examples did not enhance the creative performance of the high-creative group when compared to the low-creative group. This may result from that the level of example diversity is not high enough to activate broader problem-relevant knowledge for individuals with high creativity. That is, there are only 2 examples in the diverse or similar example condition. Future study should consider the diversity range to investigate the stimulation effect of example diversity on post-exhaustion DT performance.

4. General discussion

This study examined how example features (i.e. idea novelty and diversity) affect the example effect on idea generation after creative exhaustion. Moreover, individual differences in creativity were also considered when exploring the effect of example diversity on idea generation after creative exhaustion. Results showed that presenting high-novel examples to individuals after creative exhaustion led to high idea originality when compared to low-novel examples or no example. Results of example diversity effect revealed that diverse examples facilitated fluency and flexibility in comparison to similar examples, whereas similar examples enhanced originality after exhaustion in comparison to diverse examples. Regardless of example diversity, high-creative participants performed better in terms of idea fluency, originality, and flexibility in comparison with low-creative ones.

Experiment 1 revealed that examples with high novelty induced higher DT originality than those with low novelty and control group after creative exhaustion. This partly supported hypothesis 1. Drawing on the associative theory of creativity, more remote associations can be activated as idea generation proceeds (Mednick, 1962). With the spreading activation of semantic network in the late stage of task, novel examples as retrieval cues may be more easily integrated within the search for ideas (Nijstad & Stroebe, 2006). Previous studies showed that avoiding using common examples could enhance originality in comparison with using common examples and explained that common examples caused mental fixation on DT (George & Wiley, 2020). In this case, we suggested that since ideas with low novelty were nearly exhausted during the early state of task, presenting such examples might be less effective in stimulating more remote associations than those with high novelty. In contrast, high-novel examples might stimulate more remote associations, lead individuals to concern unexplored ideas, and boost higher originality after creative exhaustion.

Participants that were exposed to high-novel examples showed no advantage on post-exhaustion fluency and flexibility over those with medium-novel or low-novel examples. The possible reason is that individuals may have difficulty in generating a large number of new ideas after creative exhaustion even after seeing examples. That is, the equal post-exhaustion fluency and flexibility between groups may result from floor effect. In addition, the low-novel group performed better than the control group. According to the associative theory of creativity mentioned above, low-novel examples might activate close associations and relevant knowledge or even direct individuals to deeply explore

candidates within the domain of the examples. In this case, individuals are capable of generating more common or original ideas, thereby the low-novel group showed high fluency, originality, and flexibility in comparison to the control group.

In line with hypothesis 2, experiment 2 showed that post-exhaustion fluency and flexibility was higher in the diverse example condition than that in the similar example condition. This finding supported the SIAM model, which suggests diverse stimuli can activate more problem-relevant knowledge than similar stimuli (Nijstad & Stroebe, 2006). Moreover, post-exhaustion originality was lower in the diverse example condition than that in the similar example condition. As the similar example group had lower pre-exhaustion originality than the diverse example group, it still had room to improve its originality after-exhaustion. Given the max-3 scoring results did not show such difference, this might also be an effect of the average scoring used in this study. Alternatively, according to the attention allocation theory, the disparity activation of relevant knowledge domains by diverse examples will inhibit deep exploration and reduce idea originality as a result (Sio et al., 2015). Moreover, holding diverse knowledge in mind simultaneously can increase cognitive load and reduce cognitive resources in ideation (Santanen et al., 2004). As participants had only 5 min to generate ideas after seeing the examples, the similar examples may be processed more easily and deeply than the diverse ones. Therefore, the similar examples could stimulate higher idea originality than the diverse ones. However, these explanations should be taken cautiously and deserved further explorations.

Hypothesis 3 partly suggested that individuals' creativity may affect effect of example diversity on post-exhaustion DT performance because of their differences in semantic network structure (Kenett et al., 2014). The high-creative group performed better than the low-creative ones in terms of fluency, flexibility and originality after creative exhaustion, thereby Hypothesis 3 was partly supported. However, experiment 2 showed no interaction effect between Creativity Level and Example Diversity on creativity after creative exhaustion. Therefore, hypothesis 3 was not completely supported. The possible reason is that only two examples were presented in the diverse or similar example condition, and thereby the manipulated example diversity was not diverse enough. In addition, participants only had 5 min to use these examples after creative exhaustion in this study. As the associative memory of high-creative person is broader and more flexible (Kenett et al., 2014), more time may be necessary for individuals to recognize and use the examples as retrieve cues. Future research should take these possible factors into consideration as well.

To sum up, the findings about the effect of example novelty, as well as example diversity on DT performance after creative exhaustion, advanced our understanding about the role of example features in creative performance. Previous studies only investigated the difference between common (i.e., low-novel ones) and original examples (i.e., high-novel ones). The current study further divided examples into high, medium and low-novel ones. While example timing has been investigated by previous studies (Siangliulue et al., 2015; Tseng et al., 2008; Yuan et al., 2021), this study went further and examined the example effect on DT performance after creative exhaustion. These findings revealed how example novelty and diversity affected the example effect on DT performance after creative exhaustion, which help support and refine theories of examples' inspiration. It is worth noting that the high-creative and low-creative groups differed more in the post-exhaustion session than in the pre-exhaustion session. It is possible that high-creative individuals can utilize examples better than low-creative ones. This finding suggests that the role of individual differences in creativity should be considered in example inspiration studies.

This study has several limitations. Firstly, the precise mechanism underlying the stimulation effect of example novelty and diversity on post-exhaustion DT performance was not explored. Further research may use neuroscientific method or others to explain these findings based on theories such as the SIAM model. Secondly, given only two examples were presented in the diverse example condition, the manipulated example diversity might not be diverse enough. This might mask the hypothesized benefits of diverse examples for the high-creative individuals. Moreover, participants only had 5 min to recognize and use these examples in this study. Further research could test the impact of diversity range and post-exhaustion task duration in the example effect on post-exhaustion DT performance. Thirdly, creative exhaustion was identified as the moment when participants reported that they had run out of ideas in this experimental design. This led to a dilemma that the time on task (and number of ideas) before exhaustion, which was systemically associated with creativity performance, differed across participants. Therefore, further research can manipulate creative exhaustion in other ways such as presenting examples after a fixed time period, fixed number of ideas, or after idea fluency goes below a certain threshold. Lastly, a messenger app was used in the current study to collect data due to the outbreak of COVID-19. Future studies should adopt lab design to ensure task commitment and valid performance.

CRediT authorship contribution statement

H.Y., C.Y., and N.H. designed the experiment. H.Y., C.Y., and M.J. performed the study. H.Y. and K.L. intensively discussed the analytical plans and H.Y. analyzed the data. H.Y., K.L., M.J., C.Y., and N.H. wrote the manuscript. H.Y., K.L., and N.H. reviewed and edited the manuscript. H.Y. and K.L. equally contributed to the work.

Declaration of competing interest

None.

Acknowledgements

This work was supported by the Humanity and Social Science Foundation of Ministry of Education of China (NO. 21YJC190021; NO. 19YJA190010) and Philosophy and Social Science Foundation of Colleges and Universities in Jiangsu Province (NO. 2021SJA1401). We thank Qiong Wu for his help in the artwork modification. We also thank Meng Liu, Jia Xing, He Wang and Xiaoyue Jin for their help in collecting the data.

References

Althuizen, N., & Wierenga, B. (2014). Supporting creative problem solving with a case-based reasoning system. *Journal of Management Information Systems, 31*, 309–340. https://doi.org/10.2753/MIS0742-1222310112

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45*, 357–376. https://doi.org/10.1037/0022-3514.45.2.357

Baruah, J., & Paulus, P. B. (2011). Category assignment and relatedness in the group ideation. *Journal of Experimental Social Psychology, 47*, 1070–1077. https://doi.org/10.1016/j.jesp.2011.04.007

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 309–319. https://doi.org/10.1037/a0029171

Chen, Q., Beaty, R. E., Wei, D., Yang, J., Sun, J., Liu, W.Qiu, J., … (2016). Longitudinal alterations of frontoparietal and frontotemporal networks predict future creative cognitive ability. *Cerebral Cortex, 28*, 103–115. https://doi.org/10.1093/cercor/bhw353

Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology, 53*, 82–88. https://doi.org/10.1037/h0045461

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

Fink, A., Koschutnig, K., Benedek, M., Reishofer, G., Ischebeck, A., Weiss, E. M., & Ebner, F. (2012). Stimulating creativity via the exposure to other people's ideas. *Human Brain Mapping, 33*, 2603–2610. https://doi.org/10.1002/hbm.21387

George, T., & Wiley, J. (2020). Need something different? Here's what's been done: Effects of examples and task instructions on creative idea generation. *Memory & Cognition, 48*, 226–243. https://doi.org/10.3758/s13421-019-01005-4

George, T., Wiley, J., Koppel, R. H., & Storm, B. C. (2019). Constraining or constructive? The effects of examples on idea novelty. *The Journal of Creative Behavior, 53*, 396–403. https://doi.org/10.1002/jocb.178

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*, 611–625. https://doi.org/10.1111/j.2044-8295.2007.tb00467.x

Gray, C. M., McKilligan, S., Daly, S. R., Seifert, C. M., & Gonzalez, R. (2019). Using creative exhaustion to foster idea generation. *International Journal of Technology & Design Education, 29*, 177–195. https://doi.org/10.1007/s10798-017-9435-y

Guilford, J. P. (1967). *The nature of human intelligence*. NY: McGraw-Hill. https://doi.org/10.1017/9781316817049

Hao, N., Yuan, H., Hu, Y., & Grabner, R. H. (2014). Interaction effect of body position and arm posture on creative thinking. *Learning & Individual Differences, 32*, 261–265. https://doi.org/10.1016/j.lindif.2014.03.025

Hass, R. W., & Beaty, R. E. (2018). Use or consequences: Probing the cognitive difference between two measures of divergent thinking. *Frontiers in Psychology, 9*, 2327. https://doi.org/10.3389/fpsyg.2018.02327

Jonathan, D. R., Cuthbert, A. S., & Tynan, M. E. (2021). The neglect of idea diversity in creative idea generation and evaluation. *Psychology of Aesthetics, Creativity, and the Arts, 15*, 125–135. https://doi.org/10.1037/aca0000235

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience, 8*, 407. https://doi.org/10.3389/fnhum.2014.00407

Koppel, R. H., & Storm, B. C. (2014). Escaping mental fixation: Incubation and inhibition in creative problem solving. *Memory, 22*, 340–348. https://doi.org/10.1080/09658211.2013.789914

Lambert, K. G., Moss, J., & Cagan, J. (2019). A neuroimaging investigation of design ideation with and without inspirational stimuli—Understanding the meaning of near and far stimuli. *Design Studies, 60*, 1–38. https://doi.org/10.1016/j.destud.2018.07.001

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

Moss, J., Kotovsky, K., & Cagan, J. (2007). The influence of open goals on the acquisition of problem-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 876–891. https://doi.org/10.1037/0278-7393.33.5.876

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: A cognitive model of idea generation in groups. *Personality and Social Psychology Review, 10*, 186–213. https://doi.org/10.1207/s15327957pspr1003_1

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2002). Cognitive stimulation and interference in groups exposure effects in an idea generation task. *Journal of Experimental Social Psychology, 38*, 535–544. https://doi.org/10.1016/S0022-1031(02)00500-0

Perttula, M., & Sipilä, P. (2007). The idea exposure paradigm in design idea generation. *Journal of Engineering Design, 18*, 93–102. https://doi.org/10.1080/09544820600679679

Pi, Z., Hong, J., & Hu, W. (2019). Interaction of the originality of peers' ideas and students' openness to experience in predicting creativity in online collaborative groups. *British Journal of Educational Technology, 50*, 1801–1814. https://doi.org/10.1111/bjet.12671

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*, 66–75. https://doi.org/10.1080/10400419.2012.652929

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

Runco, M. A., & Mraz, W. (1992). Scoring divergent thinking tests using total ideational output and a creativity index. *Educational and Psychological Measurement, 52*, 213–221. https://doi.org/10.1177/001316449205200126

Runco, M. A., & Pritzker, S. R. (1999). *Encyclopedia of creativity*. San Diego, Calif: Academic Press.

Santanen, E. L., Briggs, R. O., & Vreede, G.-J.d. (2004). Causal relationships in creative problem solving: comparing facilitation interventions for ideation. *Journal of Management Information Systems, 20*, 167–198. https://doi.org/10.1080/07421222.2004.11045783

Siangliulue, P., Chan, J., Gajos, K. Z., & Dow, S. P. (2015). Providing timely examples improves the quantity and quality of generated ideas. In *Paper presented at the proceedings of the 2015 ACM SIGCHI conference on creativity and cognition, New York, USA*. https://doi.org/10.1145/2757226.2757230

Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *Design Studies, 39*, 70–99. https://doi.org/10.1016/j.destud.2015.04.004

Smith, S. M. (1995). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg, & J. E. Davidson (Eds.), *The nature of insight* (pp. 229–251). The MIT Press.

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3–15). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511807916.003.

Tseng, I., Moss, J., Cagan, J., & Kotovsky, K. (2008). The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies, 29*, 203–221. https://doi.org/10.1016/j.destud.2008.01.003

Wang, K., Nickerson, J., & Sakamoto, Y. (2018). Crowdsourced idea generation: The effect of exposure to an original idea. *Creativity and Innovation Management, 27*, 196–208. https://doi.org/10.1111/caim.12264

Wang, M., Hao, N., Ku, Y., Grabner, R. H., & Fink, A. (2017). Neural correlates of serial order effect in divergent thinking. *Neuropsychologia, 99*, 92–100. https://doi.org/10.1016/j.neuropsychologia.2017.03.001

Yuan, H., Lu, K., Yang, C., & Hao, N. (2021). Examples facilitate divergent thinking: The effects of timing and quality. *Consciousness and Cognition, 93*, Article 103169. https://doi.org/10.1016/j.concog.2021.103169