Title
How does comparing (dis)similar objects affect young children's creative idea generation? Exploring the role of diversity in facilitating creativity

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Koo, Chanhee
Bai, Honghong
Luo, Aoxin
et al.

Publication Date
2024

Peer reviewed

How does comparing (dis)similar objects affect young children's creative idea generation? Exploring the role of diversity in facilitating creativity

Chanhee Koo (chanheekoo98@gmail.com)
Department of Psychology & Tsinghua Laboratory of Brain and Intelligence, Tsinghua University
Beijing, China, 100084

Honghong Bai[1] (honghong.bai@ru.nl)
Behavioural Science Institute & Orthopedagogics: Learning and Development, Radboud University
Thomas Van Aquinostraat 4, 6525GD Nijmegen, the Netherlands

Aoxin Luo (al5dbh@virginia.edu)
Tsinghua Laboratory of Brain and Intelligence, Tsinghua University
Beijing, China, 100084

Stella Christie (christie@tsinghua.edu.cn)
Department of Psychology & Tsinghua Laboratory of Brain and Intelligence, Tsinghua University
Beijing, China, 100084

Abstract

This research tested and confirmed a novel hypothesis that similarity sets the ground for diversities to emerge, which then give rise to creativity. Adopting an experimental design, we recruited 66 typically-developing Chinese children (M = 6.04 years, SD = .28). First, in a Comparison task, these children were randomly assigned to name differences between two objects that were either highly similar (high-similarity condition) or dissimilar (low-similarity condition). Next, all children completed a divergent thinking task and received scores on fluency, originality, and usefulness. Results of t-tests showed that children of high-similarity condition reported both more surface (pertaining to perceptual features) and deep (pertaining to structural features) alignable differences and have on average a higher originality score, compared to children of low-similarity condition. Mediation analysis results further showed that the number of deep alignable differences mediated the effect of condition on children's originality scores. This confirmed our expectation that the high similarity between objects facilitated children to generate more deep alignable differences, which subsequently facilitated children to generate more original ideas.

Keywords: creativity, analogical reasoning, comparison, similarity, diversity

Introduction

Creativity, a central pillar of human cognitive prowess, is a complex capacity that is unique to our species and thus, has attracted abundant research interest in the recent decades (Beghetto et al., 2001; Mejia et al., 2021; Simonton, 2000). While controversies exist, researchers generally agree that creativity pertains to the generation of products or ideas that are (potentially) original and effective (Corazza & Lubart, 2020; Runco & Jaeger, 2012). Complementing this outcome-focused definition, recent theories emphasize the dynamic nature of creativity (Corazza, 2016) and call for more empirical studies respecting how creativity emerges.

A common belief revolving around this topic is that diversity gives rise to creativity. In education, for instance, teachers often encourage students to "think differently" (Mullet et al., 2016) and/or to "think outside the box" (Akyıldız & Çelik, 2018) as a way to foster creativity. In addition, this notion is particularly true when we take into account that substantial research has adopted divergent thinking tasks to measure creativity (Reiter-Palmon et al., 2019). In these tasks, participants are often prompted with a cue and then asked to think in multiple, diverse directions (e.g., "What can you use a brick for?" in the Alternative Uses Task [AUT], developed based on Guilford, 1967). Indeed, empirical studies showed that measures of divergent thinking in (early) childhood predict people's creative achievement later (Plucker, 1999; Runco et al., 2010). For example, Torrance (1980, 1981) followed a group of primary school students starting from 1958, measuring their intelligence, academic achievement, and creativity using a battery of standardized tests every year, for six years. In this phase, the *Torrance Tests of Creative Thinking (TTCT)* were developed and used to measure children's creativity through eliciting their divergent thinking (Acar & Runco, 2012; Baer, 2011). After 22 years, these then-elementary school students were again contacted and got assessed on their creative achievement using the *Creative Style of Life Achievements* questionnaire, which tackled their everyday personal creative activities and achievements in detail and involved experts to rate the level of creativity based on these activities and achievements. When analyzing these data together, Torrance found that the TTCT scores obtained early on during elementary school were strongly correlated with the same participants' creative achievement in adulthood, even after more than two decades.

Also the line of studies associating individuals' multicultural or multilingual experiences—as sources of diversity—with their level of creativity lend support to this

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

belief. For example, Kharkhurin (2008, 2009) found that proficient bilinguals outperform monolinguals in generating and processing unrelated ideas and in extracting novel, unique ideas in divergent thinking tasks. A survey-based study showed that individuals who can speak multiple languages and/or have been exposed to multiple cultures through parents' cultural roots or by living in different places were able to generate more unconventional, novel ideas compared to their counterparts (Leung et al., 2008). More robustly, a laboratory experimental study revealed that even being shortly exposed to multicultural stimuli can enhance people's creativity (Leung & Chiu, 2010). In this study, European-American participants were first exposed to either a single unfamiliar culture (images of cultural elements of China) or exposed to two cultures, one familiar and one unfamiliar (images of cultural elements of the United States and China). They were then asked to write stories for Turkish children—a totally different culture to rule out the influence of American and Chinese cultures on the creation of stories. The stories written by participants from the multiple culture-exposure group were rated higher on novelty and uniqueness than the stories written by participants from the single culture-exposure group.

When shifting from individuals to the team settings, although positive evidence is present (e.g., Jang, 2017; Sung & Choi, 2019), it should be highlighted that more diversity does not always lead to more creativity. Bodla et al. (2018) associated the heterogeneity of demographic attributes (e.g., gender, age, ethnicity) of students affiliated to various research lab teams, as a proximal measure on the level of diversity, with the level of creativity of those teams (as rated by the lab PIs). They found that the level of intra-team diversity was negatively associated with the level of team creativity. From a different angle, Lee and Park (2020) examined how the level of diversity in team agreeableness, namely to what extent team members differ in their disposition of being trusting, compliant, caring, and gentle, was related to the level of team creativity. They found that teams with a higher level of diversity, thus mixing people of higher agreeableness with people of lower agreeableness, showed poorer performance on the team creativity task.

What is responsible for the discrepancy in the findings between studies focusing on individuals and teams? A plausible inference is that diversity is only conducive to creativity when a common ground is established. For individuals with multicultural or multilingual experiences, they are themselves the common ground—they tend to constantly align and compare different cultures (Cheng et al., 2011). In this way, they can benefit from the diversities through two mechanisms. First, they can draw inspiration from two or multiple sets of ideas, concepts, mental representations of objects, and behavioral scripts. Critically, they might even see more nuanced varieties about one thing (Leung et al., 2008) and more relations between seemingly disparate things (Cheng et al., 2008) that are likely to be overlooked in a single culture. Second, they are also more likely confronted with and 'forced' to reconcile disparate ideas and concepts along with two or more cultures. As a byproduct of this process, these people might also get trained in establishing new relations and meanings, and as summarized by van Dijk et al. (2018), get enhanced on cognitive functions such as sustained attention, cognitive flexibility, and working memory which are crucial for creative exploration and problem-solving.

In team settings, in contrast, no common ground is readily available, and diversity itself is naturally impeding the creation of a common ground. Related to studies mentioned above, for instance, Bodla et al. (2018) inferred that the diversity within the team has incited social categorization, discouraging team members to interact and to share and to take in each other's thoughts and ideas. And in Lee and Park (2020), one explanation was that a larger variance in agreeableness among team members might have introduced more task-related as well as interpersonal conflicts. In these scenarios, the members are basically left alone and thus, cannot benefit from diversities. What is worse is that people might even consciously choose to not be creative as they might feel being judged by others.

In fact, research from neighboring domains/fields also suggest that diversity and similarity—a way of measuring the "common ground" function together to give rise to creativity. First, research on analogical and relational reasoning revealed that similarity leads people to detect more diversities and enhances their thinking in establishing new, non-obvious links between different things (e.g., Gelman et al., 2009; Gentner & Markman, 1994; Markman & Gentner, 1993; Sagi et al., 2012). In those studies, participants are presented with and requested to name differences between pairs of stimuli that are either highly similar or dissimilar and asked to name the differences between the two objects involved in the stimuli-pairs. A particular robust finding from these studies is that participants consistently named more *alignable differences* between two similar objects than between two dissimilar objects. Alignable differences are the differences that can distinguish objects on a common dimension (Markman & Gentner, 1993). For instance, 'a kitten is younger than a cat' is an alignable difference because it distinguishes 'kitten' from 'cat' on the dimension of age.

A plausible conjecture is that the large set of obvious features shared between similar objects (kitten vs. cat) might have made it easier for participants and enhanced their tendency to align two objects on comparable dimensions (the age). In contrast, for objects that are radically dissimilar from each other (kitten vs. newspaper), it is rather counterintuitive for participants to align them on any dimension. Instead, the comparison would gravitate towards the direction of "one has it but the other one does not" (e.g., "the kitten has hair but the newspaper does not"). On a deeper level, the alignability between similar objects can also 'force' participants to pay attention to nuanced variations as embedded in objects and push them to even go beyond perceptual features and establish new, non-obvious (oftentimes abstract) dimensions for comparison (e.g., "kittens are more energetic, while cats are more calm" as for the behavior of a kitten vs. cat). This could, however, be difficult to achieve for comparing dissimilar object-pairs, for which one's attention is likely occupied by obvious perceptual features that are immediately accessible.

According to the newest evidence, the precise ability of looking through the perceptual features of objects and establishing the deeper link between two objects seems to be critical for the emergence of novel ideas (Bai, Leseman et al., 2021; Bai, Mulder et al., 2021). In their studies, Bai et al. asked 4 to 6 years old children to think aloud while performing the AUT task. Thus, each time children generated an idea, they were prompted to explain how they came up with this idea. By analyzing children's verbal explanations, it was found that the most effective strategy for generating truly novel ideas involves the mental deconstruction of holistic object representations and/or reconstruction of a new whole based on the break-down parts of objects (with other materials). They explained that the representation of objects as complete units may cement them to their traditional functionalities, thereby hindering children to generate unconventional ideas. In contrast, the mental practice of dissecting objects into constituent parts and properties breaks the original relational structures (e.g., an umbrella becomes a stick and a cloth cover). This facilitated the thinking of other possible relations between these broken-down parts and properties, thus opening the possibilities to more uncommon, novel ideas.

Taken together, similarity is likely the ground for diversities pertaining to different dimensions that relate different objects, which then leads to more creativity. The current study is set to test this route with a laboratory experiment, with children aged 5 to 6 years. Children were first tested with a Comparison task in which they were asked to name differences between two objects, under either high-similarity or low-similarity conditions. As outlined above, we expected children under the high-similarity condition to find more alignable differences, especially differences pertaining to relational dimensions that go beyond perceptual features of the objects ("deep alignable differences," which are to be defined below). All children were then asked to complete the AUT task in which they needed to generate as many unusual uses as possible for everyday objects. We expected that children would cast their attention and think more on the structural, non-perceptual features of objects after the Comparison task (associated with deep alignable differences), and this tendency would carry over and influence children's performance on the subsequent AUT task. Thus, we expected that children of the high-similarity condition would generate more original ideas in this task, compared to children of the low-similarity condition. The current study is novel in testing a novel hypothesis emerged from synthesizing different lines of literature, which may elucidate how similarity and diversity together give rise to creativity.

Method

Design, Participants, and Procedure

We adopted a between-group experimental design. A total of 66 typically-developing Chinese children aged 5 to 6 years (M = 6.04 years, SD = .28) were recruited to take part in the experiment, who were randomly assigned to either the high-similarity condition (N=33, including 13 girls and 20 boys; M = 6.07 years, SD = .25) or the low-similarity condition (N=33, including 11 girls and 22 boys; M = 6.00 years, SD = .30). Children of both conditions were requested to first complete a comparison task and then a divergent thinking task. The comparison task consisted of stimulus-pairs of highly similar images in the high similarity condition and stimulus-pairs of dissimilar images in the low similarity condition. The divergent thinking task was the same in both conditions.

The participants were recruited from the subject pool of the laboratory that hosted the study, located in the metropolis, capital city of Beijing. We published the project descriptions and the invitation of participation on the platform, through which parents voluntarily signed their children up for participating in the study. The demographic data showed that the majority (93.9%) of participating children were from families that reside in urban areas (thus mainly from the middle class).

The testing took place in the laboratory. Parents signed a consent form prior to the testing. Children were then tested alone in a room. The testing was live streamed on a screen outside the room so that parents could follow. To start, the experimenter welcomed the child and introduced the whole experiment. Then, the experimenter guided children to first finish the comparison and then the divergent thinking task. In between the two tasks, children had a one-minute break to play an interactive, fun card game with the experimenter, preventing their attention being overly taxed with two tasks in a row. All tasks were conducted in Mandarin Chinese.

Materials and Measures

Comparison Task. We used the comparison paradigm developed by Gelman and colleagues (2009) with minor adaptations. Children were presented with stimulus-pairs, made of pictures of animals and/or everyday objects, and asked to name the differences between the two stimuli.

Materials. To warm-up, children were asked to locate the obvious differences between three image-pairs, one difference per pair (e.g., a circle vs. the same circle wearing a hat). For the test trials, a total of 28 images (printed on white paper and laminated) consisting of 14 animals and 14 objects were used. In the high-similarity condition, animals were paired with animals, and objects were paired with objects, resulting in 14 similar image pairs (seven of animal and seven of object). In the low-similarity condition, the same images were used, but the pairs were made by each of them having one animal and one object, resulting in 14 dissimilar animal-object image pairs. The presenting order of image pairs was counterbalanced within each condition.

Test procedures. Based on Gelman et al. (2009), the current study employed generic language (e.g., "think about dogs and cats", instead of "think about this dog and this cat") in giving the instructions. This was to activate the full representation of stimuli in children's mind and their everyday experience with these stimuli.

Children were first introduced to the task and asked to finish three warm-up trials. In this step, children received corrective feedback. Then the experimenter began formal trials, saying "*I want you to think about [one stimulus, e.g., "dogs"] and [another stimulus, e.g., "cats"]. Can you tell me if and how dogs and cats are different?*" while presenting similar or dissimilar image pairs to children.

Coding and scoring. Audio recordings of test sessions were transcribed, first by an auto-transcribing software (https://rec.sogou.com/voice) and then checked and corrected by the first author to ensure that children's explanations of ideas were precisely and verbatim transcribed. The explanations were divided into episodes, with each episode spanning one or multiple sentences that expresses a continuous thought. As observed, children often named more than one difference within one episode. Thus, episodes were further segregated into responses, with each response relating to only ONE difference. Finally, each of the responses was coded as an alignable or a non-alignable difference. An auxiliary category "other" was included for coding responses not relevant to a difference such as simply naming a feature of a stimulus. The first and the third author together coded all transcripts. Based on six transcripts that were coded by both, Cohen's Kappa was .79, indicating substantial inter-rater agreement.

Alignable differences were defined as differences that either occur along a continuous dimension or are binary characteristics that occupy corresponding roles in people's representations of the stimulus items in accordance with McGill (2002). To capture nuanced variations, alignable differences were further coded into three sub-categories:

- *Surface alignable differences.* This type of difference pertains to perceptual properties directly observable from the images. Differences based on the negation of a characteristic of one comparing stimulus can be considered as an alignable difference, if two coders agree that the absence of the characteristic is represented as an explicit property which reveals about the stimulus (Gentner & Markman, 1994).
- *Deep alignable differences.* This type of difference pertains to properties that are not directly observable from the perceptual examination of the comparing stimuli, such as behaviors, functions, traits, internal parts, mental states, and taxonomic judgments, which are causally connected to each other and other deep features, as based on Gelman et al. (2009). Differences based on the negation of a characteristic of one comparing stimulus can be considered as an alignable difference if two coders agree that the absence of the characteristic is represented as an explicit property which is revealing about the stimulus.

Non-alignable differences pertain to characteristics of one comparing stimulus which cannot be matched up with characteristics in the other stimulus. An exception to this rule was when two coders agreed that people generally represent the absence of a property explicitly for a given stimulus (e.g., "a whale can swim, a camel cannot swim"), this difference would be coded as an *alignable difference*.

Alternative Uses Task. We used AUT (Guilford, 1967) to measure divergent thinking. Children were asked to give as many unusual uses as possible for four everyday objects.

Test materials and procedures. We followed Bai et al. (2021) on setting up the testing with some adaptations. Children were first introduced to the task ("*Let's begin the second game! In this game, I'm going to show you pictures of some common objects. When you see the pictures, try to think about what the object can be used for. Try to think of as many unusual and different uses as you can, preferably ones that no one else would think of. Let's practice first*"). To warm up, children were then presented with a real toothbrush and modeled by the experimenter on generating three unusual uses of it ("*Look, this is a toothbrush. We usually use toothbrushes to brush our teeth or brush things to clean them. But actually, toothbrushes can be used for many other things. If I were to think of something that no one else can think of, for example, there are a lot of bristles on the toothbrush, so if I feel itchy on my back but I can't reach it with my hand, I can use the toothbrush to scratch it. Also, the body of the toothbrush is long and hard, so I can stick it in a pile of dirt and use it as a flag. Look, the handle of the toothbrush has soft plastic, so you can use it as a teething toy for your dog! Now it's your turn. Think about what else you can use your toothbrush for. Tell me whatever you think of.*") The example uses were given to make the purpose of the task clear to children—generating uses that were *unusual* and not commonly associated with the object. In the formal test trials, children were presented a total of four images, each with an everyday object on it (printed on white paper and laminated). Each time an object was presented, children were prompted to generate original ideas (e.g. "*This is a pencil. Please think about what the pencil can be used for. Remember, try to think of unusual and different uses, preferably ones that no one else would think of*"). To reduce the cognitive load while strengthening the priming effect as conditioned by the comparison task, we selected images of a pencil, a fork, a necklace, and a bowl that were also used in the Comparison Task as stimuli for the AUT task. The selection of stimuli additionally ensured that the objects can afford various uses and are familiar to children. The AUT task was conducted verbally and audio-recorded, and the time limit was 3 minutes per object.

Coding and scoring. We first verbatim transcribed the audios. Then, each transcript was separated into the child's and the experimenter's episodes of speech, and the *uses* generated by the child were extracted for each object. We scored the originality and the usefulness of the uses.

- *Originality* refers to the extent to which the generated idea is surprising, novel, and unique while considering participants' age, on a 7-point Likert scale (from 1 ="*not original at all*" to 7= "*very original*".

- *Usefulness* refers to the extent to which the generated idea is imaginable, feasible, and helpful while considering participants' age, on a 7-point Likert scale (from 1="*not useful at all*" to 7="*very useful*").

The first and the third author together rated all responses on originality and usefulness. Based on both raters' scoring of responses given by six children, the intraclass correlation coefficients (ICC) were .90 for originality (excellent consistency) and .79 for usefulness (good consistency).

Finally, individuals' *fluency* scores were calculated by summing up the number of uses generated across objects, and their *originality and usefulness* scores were calculated by summing up the originality and usefulness scores of all responses and dividing them by individuals' fluency scores.

Results

All analyses were conducted in R. Two children's data were excluded from further analyses due to technical issues.

Descriptive and Correlational Analyses

Tables 1 and 2 present the descriptive statistics and correlations of alignable and non-alignable differences generated during the Comparison Task and the fluency, originality, and usefulness measures from the AUT.

Table 1: Descriptive statistics (*N* = 64).

Variables	*M* ± *SD*	Range (Min-Max)
Comparison Task Measures (# of responses):		
Surface Alignable Differences	26.89 ± 20.99	0.00 - 80.00
Deep Alignable Differences	10.77 ± 9.14	0.00 - 35.00
Total Alignable Differences	37.66 ± 26.14	1.00 - 105.00
Others-Non-alignable Differences	27.86 ± 23.53	0.00 - 111.00
AUT Measures (mean scores across stimuli):		
Fluency	20.88 ± 8.56	6.00 - 43.00
Originality	3.56 ± .89	1.73 - 5.53
Usefulness	5.08 ± .93	2.88 - 6.50

Table 2: Correlational analysis (*N* = 64).

Variables	2	3	4	5	6	7
1. Surface Alignable Differences	.41***	.95***	.62***	.28*	.31*	.12
2. Deep Alignable Differences		.68***	.44***	.08	.53***	-.01
3. Total Alignable Differences			.65***	.25*	.44***	.09
4. Others-Non-alignable Differences				.03	-.31*	-.13
5. Fluency					.26*	-.15
6. Originality						0.38**
7. Usefulness						

*$p < .05$. **$p < .01$. ***$p < .001$.

Effects of Condition on Task Performance

Independent sample t-tests were conducted to examine the effect of condition on children's task performance.

The Comparison Task. The results showed a significant group effect on the total number of alignable differences $t(62)=7.39$, $p<.001$, 95% CI=[26.12, 45.51], showing that children of the high-similarity group generated more alignable differences (M=56.85, SD=20.88) than children of the low-similarity group (M=21.03, SD=17.66). Specifically, children of the high-similarity group generated significantly (1) more surface alignable differences (M=40.76, SD=18.86) than children of the low-similarity group (M=14.23, SD=13.71), $t(62)=6.40$, $p<.001$, 95% CI=[18.25, 34.82] and (2) more deep alignable differences (M=15.61, SD=7.77) than children of the low-similarity group (M=6.61, SD=8.05), with $t(62)=4.55$, $p<.001$, 95% CI=[5.04, 12.95].

In contrast, children of the low-similarity group generated more non-alignable differences (M=42.71, SD=24.49) than children of the high-similarity group (M=12.85, SD=8.57), for which $t(62)=-6.59$, $p<.001$, 95% CI=[-38.92, -20.80].

AUT. There was a group effect on children's originality scores (which was controlled for fluency in deriving the scores), $t(62)=3.59$, $p<.001$, 95% CI=[0.35, 1.21], showing that children of the high-similarity group generated, on average, more original ideas (M=3.94, SD=0.84) than children of the low-similarity group (M=3.16, SD=0.90). Nevertheless, there was no group effect on children's fluency, $t(62)=1.00$, $p=.32$, 95% CI=[-2.12, 6.37], and usefulness $t(62)=.24$, $p=.81$, 95%, CI= [-0.41, 0.52].

Relation of the Comparison Task and the AUT

We also explored whether the relationship between the level of similarity with creativity was established through facilitating the generation of alignable differences. We carried out two times of mediation analysis (Baron & Kenny, 1986), one with the number of surface-alignable differences as the mediator and the other one with the number of deep-alignable differences as the mediator.

Deep alignable differences. We first established a baseline model with originality scores regressed on the condition. The results of this model showed a significant effect of condition (high similarity=0 & low similarity=1) on children's originality scores, $\beta=-0.78$, $p<.001$, indicating children of the high-similarity group, on average, received higher originality scores compared to children of the high-similarity group. Next, it was established that the condition significantly predicted the number of deep alignable differences, and the number of deep alignable differences significantly predicted originality scores when controlling for the condition. Finally, when the generation of deep-alignable differences was controlled for, the direct effect of similarity on originality was notably reduced and became insignificant ($p=.08$). These results, together with a Sobel test (z=-2.53, $p<.05$), confirmed that the number of deep-alignable differences mediated the effect of condition on children's originality scores.

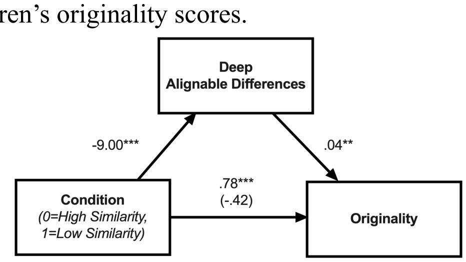

Figure 1. Deep-alignable difference mediated the effect of condition (similar vs. dissimilar) on children's originality.

Surface alignable differences. The number of surface alignable differences seemed to also partially mediate the effect of condition on children's originality, but the results were not fully robust and hence not reported.

Discussion

The impetus of the current study was to investigate how comparison of high-similarity and low-similarity image pairs affects young children's creative idea generation. Specifically, we sought out to answer whether comparing similar objects can prime children to locate their attention and thinking more on objects' structural, non-perceptual features, which may subsequently enhance their creativity. Theoretically, this study attempted to test a novel hypothesis which points to that similarity gives rise to (relational) diversities as to give rise to creativity.

One main finding from our analyses was that, when presented with high-similarity stimulus pairs, children indeed reported more alignable differences, both surface and deep. This is consistent with past research on adults wherein it is quite robustly established that similar pairs nearly always prompt participants to find more alignable differences (Gentner & Markman, 1994; Markman & Gentner, 1996; Sagi et al., 2012). In addition, also in one study wherein children were included as participants, a distinction was made between surface and deep alignable differences, just like the current study. Children generated more of both surface and deep alignable differences for high-similarity pairs compared to for low-similarity pairs (Gelman et al., 2009). This further strengthens the validity of this finding, indicating that comparing similar things is indeed an effective approach to facilitate people to pay more attention and thinking about diversities pertaining to structural features that relate two stimuli. It is particularly interesting to confirm that this approach also works well with children as young as 5 or 6 years of age, who have a rather limited knowledge base about structural, non-perceptual features of objects. From this point, a practical implication might point to applying the comparison paradigm to supporting children (especially those with special needs) to learn and to expand their knowledge about the structures and relations.

Pertaining to the core research question of whether children were more creative in the high-similarity condition, our expectation was also confirmed, as reflected in the group effect in children's originality scores. Compared to originality that is usually seen as a core feature of creativity (J. K. Smith & Smith, 2017; Runco & Jaeger, 2012; Simonton, 2012; Weisberg, 2015), fluency is mostly regarded as facilitating originality (Forthmann et al., 2018; Nijstad et al., 2010; Silvia et al., 2008). In this regard, having no group effect on children's fluency scores does not object to the effectiveness of comparing similar objects on facilitating children's creativity. On the contrary, it even points out that by comparison, children's original thinking can be more directly enhanced, thus not through the route of thinking of many ideas to increase the probability of generating original ideas.

In parallel with originality, usefulness is generally considered as the other essential criteria of creativity (Runco & Jaeger, 2012). Yet, it is generally harder for people to pursue originality than usefulness. In this regard, when given a creative task, people generally tend to put more mental effort into the originality aspect. This could be particularly the case when participants are especially instructed to do so, as it was also done in the current study (e.g., we stressed more on generating uses that are unusual). This might partially explain the absence of a group effect on the usefulness scores. On the other hand, children aged 5 to 6 years are often not fully proficient in verbally explaining their ideas, especially regarding novel ideas. This might also partially explain the lack of effect on this measure, which likely also weakens the robustness of using usefulness as a measure of children's creativity. Future research is warranted to find a better way in studying the usefulness aspect of creativity.

Another important finding of the current study was that the number of deep alignable differences indeed mediated the effect of condition on children's originality scores. This is in line with our expectation that, by comparing similar objects, children might learn to cast their attention more on the structural, non-perceptual features that link two objects, and this tendency would carry over and lead children to come up with more original ideas in the subsequent AUT task. In the AUT, children needed to come up with unusual uses of an object (e.g., an umbrella), which is presented in its usual, structured form. Past research has shown that breaking down the structured form of an object is a critical strategy for participants to think outside of the box and come up with novel uses (e.g., Bai, Mulder, et al., 2021; Gilhooly et al., 2007). In the current study, by establishing deep alignable differences between similar comparing objects, children might gradually learn to predominantly pay their attention to the non-perceptual features of an object, which opens possibilities for linking one object to another object. In this way, the whole, original structure of an object might be loosened so that single features that constitute this structure get more evident, thus inspiring the generation of original ideas. This finding provides first evidence, to the best of our knowledge, for elucidating how similarity and diversity together give rise to creativity. Future research can further validate this finding by embedding similarity into creativity tasks.

By reviewing literature from both the creativity and the analogical reasoning fields, this research points out a research gap related to our common belief in "diversity gives rise to creativity". We specifically tested and confirmed the novel hypothesis that similarity sets the ground for diversities pertaining to structural, non-perceptual features of objects to emerge, which then give rise to creativity. This contributes to unraveling the emergence mechanisms of creativity and the overlapping cognitive basis between analogical reasoning and creativity.

References

Acar, S., & Runco, M. A. (2012). Creative abilities: Divergent thinking. In M. D. Mumford (Ed.), *Handbook of Organizational Creativity* (pp. 115–139). Elsevier.

Akyıldız, S., & Çelik, V. (2020). Thinking outside the box: Turkish EFL teachers' perceptions of creativity. *Thinking Skills and Creativity*, *36*, 100649.

Baer, J. (2011). How divergent thinking tests mislead us: Are the Torrance Tests still relevant in the 21st century? The Division 10 debate. *Psychology of Aesthetics, Creativity, and the Arts, 5*(4), 309–313.

Bai, H., Leseman, P. P. M., Moerbeek, M., Kroesbergen, E. H., & Mulder, H. (2021). Serial order effect in divergent thinking in five-to six-year-olds: Individual differences as related to executive functions. *Journal of Intelligence*, *9*(2), 20.

Bai, H., Mulder, H., Moerbeek, M., Kroesbergen, E.H., & Leseman, P.P. (2021). Divergent thinking in four-year-old children: An analysis of thinking processes in performing the Alternative Uses Task. *Thinking Skills and Creativity*, *40*, 100814.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Beghetto, R. A., Plucker, J. A., & MaKinster, J. G. (2001). Who studies creativity and how do we know? *Creativity Research Journal, 13*(3-4), 351–357.Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36*(6), 1019-1
050.

Bodla, A.A., Tang, N., Jiang, W., & Tian, L. (2016). Diversity and creativity in cross-national teams: The role of team knowledge sharing and inclusive climate. *Journal of Management & Organization, 24*, 711 - 729.

Cheng, C.‐Y., Leung, A. K.‐y., & Wu, T.‐Y. (2011). Going beyond the multicultural experience—creativity link: The mediating role of emotions. *Journal of Social Issues, 67*(4), 806–824.

Cheng, C. Y., Sanchez-Burks, J., & Lee, F. (2008). Connecting the dots within: creative performance and identity integration. *Psychological science*, *19*(11), 1178–1184.

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

Forthmann, B., Szardenings, C., & Holling, H. (2018). Understanding the confounding effect of fluency in divergent thinking scores: Revisiting average scores to quantify artifactual correlation. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 94–112.

Gelman, S. A., Raman, L., & Gentner, D. (2009). Effects of language and similarity on comparison processing. *Language Learning and Development: The Official Journal of the Society for Language Development, 5*(3), 147–171.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155–170.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science, 5*(3), 152–158.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513.

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.

Guilford, J. P. 1967. *The Nature of Human Intelligence*. New York: McGraw-Hill Book Co.

Jang (2017) Cultural brokerage and creative performance in multicultural teams. *Organization Science 28*(6):993-1009.

Jellen, H. G., & Urban, K. K. (1986). The TCT-DP (Test for Creative Thinking-Drawing Production): An instrument that can be applied to most age and ability groups. *Creative Child & Adult Quarterly, 11*(3), 138–155.

Kharkhurin, A. V. (2008). The effects of linguistic proficiency, age of second language acquisition, and length of exposure to a new cultural environment on bilinguals' divergent thinking. *Bilingualism: Language and Cognition, 11*(2), 225–243.

Kharkhurin, A. V. (2009). The role of bilingualism in creative performance on divergent thinking and invented alien creatures tests. *Journal of Creative Behavior, 43*(1), 59–71.

Lee, S. T. H., & Park, G. (2020). Does diversity in team members' agreeableness benefit creative teams? *Journal of Research in Personality, 85,* Article 103932.

Leung, A. K., & Chiu, C. (2010). Multicultural experience, idea receptiveness, and creativity. *Journal of Cross-Cultural Psychology, 41*(5-6), 723–741.

Leung, A. K., Maddux, W. W., Galinsky, A. D., & Chiu, C. Y. (2008). Multicultural experience enhances creativity: the when and how. *The American Psychologist*, *63*(3), 169–181.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*(4), 431–467.

Mejia, C., D'Ippolito, B., & Kajikawa, Y. (2021). Major and recent trends in creativity research: An overview of the field with the aid of computational methods. *Creativity and Innovation Management, 30*(3), 475–497.

Mullet, D. R., Willerson, A., Lamb, K. N., & Kettler, T. (2016). Examining teacher perceptions of creativity: A systematic review of the literature. *Thinking Skills and Creativity, 21,* 9–30.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: young children's use of comparison in category learning. *Journal of Experimental Psychology. General, 131*(1), 5–15.

Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model:

Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, *21*(1), 34–77.

Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal, 12*(2), 103–114.

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

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.

Runco, M. A. (2004). Creativity. *Annual Review of Psychology, 55*, 657–687.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*(1), 66–75.

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

Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance tests of creative thinking as predictors of personal and public achievement: A fifty-year follow-up. *Creativity Research Journal, 22*(4), 361–368.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 68–85.

Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. *American Psychologist, 55*(1), 151–158.

Smith, J. K., & Smith, L. F. (2017). The 1.5 criterion model of creativity: Where less is more, more or less. *Journal of Creative Behavior*, *51*(4), 281–284.

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist, 51*(7), 677–688.

Sung, S. Y., & Choi, J. N. (2019). Effects of diversity on knowledge sharing and creativity of work teams: Status differential among members as a facilitator. *Human Performance, 32*(3-4), 145–164.

Torrance, E. P. (1981). Predicting the creativity of elementary school children (1958-80) —and the teacher who "made a difference." *Gifted Child Quarterly*, *25*(2), 55-62.

Torrance, E. P., & Horng, R. y. (1980). Creativity and style of learning and thinking characteristics of adaptors and innovators. *Creative Child & Adult Quarterly, 5*(2), 80–85.

van Dijk, M., Kroesbergen, E. H., Blom, E., & Leseman, P.P. (2018). Bilingualism and creativity: Towards a situated cognition approach. *Journal of Creative Behavior, 53*, 178-188.

Walia, C. (2019). A dynamic definition of creativity. *Creativity Research Journal, 31*(3), 237–247.

Weisberg, R. W. (2015). On the usefulness of "value" in the definition of creativity. *Creativity Research Journal*, *27*(2), 111–124.