A novel coding scheme for assessing responses in divergent thinking: An embodied approach

Heath Matheson, **Yoed N. Kenett**

Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

In the present study we devised a novel coding scheme for responses generated in a divergent thinking task. Based on considerations from behavioural and neurocognitive research from an embodied perspective, our scheme aims to capture dimensions of simulations of action or the body. In an exploratory investigation, we applied our novel coding scheme to analyze responses from a previously published dataset of divergent thinking responses. We show that a) these dimensions are reliably coded by naïve raters, and that b) individual differences in creativity influences the way in which different dimensions are used over time. Overall, our results provide new hypotheses about the generation of creative response in the divergent thinking task and should serve to characterize the cognitive strategies used in creative endeavors.

Keywords

embodied cognition; divergent thinking; creativity; cognitive strategies

Creative ability is defined by the ability to generate novel and appropriate new ideas (Runco & Jaeger, 2012; Simonton, 2016; Sowden, Pringle, & Gabora, 2014). Commonly applied creativity tasks are divergent thinking (DT) tasks which require participants to generate responses to open-ended questions (Acar & Runco, 2019; Runco & Acar, 2012). For instance, in the alternative uses task, participants are required to generate alternative, novel, and creative uses to common objects (Torrance, 1966). When shown the image of a shoe a participant in this task may suggest that the sole of the shoe could be used to hammer a nail into the wall. Typically, these responses are then measured on various dimensions such as novelty, uniqueness, or appropriateness (see Vartanian et al., 2019 for a review). While DT tasks have been applied in creativity research for decades, research has solely focused on the outputs in this task and very little is known about the cognitive strategies that people use when required to generate such creative, alternative uses (Hennessey & Amabile, 2010; Runco & Acar, 2012). The limited research examining potential strategies suggests that individuals examine potential action related uses when generating responses (Gilhooly, Fioratou, Anthony, & Wynn, 2007). Given this, the aim of the current study is to expand the extant research on potential cognitive strategies using an embodied cognition framework.

Corresponding author: Heath E. Matheson, University of Northern British Columbia, Department of Psychology, Prince George, BC, Canada, heath.matheson@unbc.ca, Phone number: 12509605002.

Currently, to the best of our knowledge, only a small number of attempts have been made to characterize the types of strategies participants use by analyzing the content of the verbal responses in DT. First, Gilhooly et al. (2007) showed that participants tend to focus on the properties of objects, imagine disassembling them, and produce broad uses that may serve a creative purpose (e.g. a *shoe* as *art*). This finding is supported by research that has explicitly instructed participants to disassemble the objects. For instance, studies have shown that instructions to use the disassembly strategy do indeed increase people's creative output on DT (Nusbaum & Silvia, 2011; Wilken, Forthmann, & Holling, 2019). This also occurs for other tasks, including figural DT (Forthmann et al., 2016). Similarly, an early exploratory study used a think aloud protocol during a DT and showed that responses could be reliably sorted into categories related to structuring the problem, performing a memory search, and evaluating the outcome (Khandwalla, 1993). Importantly, this study suggested that one of the most effective strategies was 'probing' or 'elaborating' on possible solutions, and from their examples given, this appeared often to involve the disassembly strategy. These findings support the idea that that the disassembly cognitive strategy successfully contributes to creative performance.

Taking a different approach, Chrysikou et al. (2016) developed a coding scheme that assessed whether participants relied on the concrete perceptual attributes of objects in producing their DT responses and showed that participants were more likely to rely on concrete properties when generating responses to words vs. pictures. More current research has focused on whether participants rely on retrieving known 'uncommon' uses from memory (e.g. *lipstick* as a *writing tool*) or whether they are generating truly novel uses (Silvia, Nusbaum, & Beaty, 2017). In children, Harrington, Block, and Block (1983) coded five-year olds' responses into a number of categories, including whether the response described a functional outcome (e.g. deliver it) or a non-functional one (e.g. stomp it) and whether the response was a common or uncommon one; the authors showed that the number of uncommon uses children gave at age 5 predicted creativity (reported by their teachers) at 11 years. Thus, these studies illustrate how alternative uses appear to rely on the production of concrete, action-related uses of objects, as one strategy that promotes performance.

These strategies may be understood within the framework of grounded or embodied cognition (Barsalou, 2008). One general proposal within the embodied cognition framework is that simulations (i.e. reactivations) of modality-related information (e.g. visual, auditory, motor) underlie our ability to identify objects and reason about possible future outcomes and situations. Stemming from this framework is the hypothesis that generating uncommon uses relies on simulating the possible actions and outcomes associated with a creative use; specifically, suggesting that a shoe can be used as a hammer requires activating simulations of the visual (i.e. aspects of shape), somatosensory (i.e. aspects of material), and motor (i.e. aspects of the bodily motions involved and the proprioceptive consequences of those motions) possibilities afforded by the shoe. Importantly, previous theorists have speculated about the role of mental imagery in creative tasks (Gilhooly et al., 2007) and research has shown that, in general, the use of mental imagery predicts creativity performance (LeBoutillier & Marks, 2003). Given the similarities between the imagery and the embodied cognitive literature (Pearson, Naselaris, Holmes, & Kosslyn, 2015), these

findings provide preliminary evidence of the usefulness of the embodied perspective in characterizing divergent thinking strategies.

To date, only a few studies have investigated creativity from an embodiment perspective (Friedman & Förster, 2000, 2002; Goldstein, Revivo, Kreitler, & Metuki, 2010; Leung et al., 2012; Oppezzo & Schwartz, 2014; Vohs, Redden, & Rahinel, 2013). However, these studies adopted broad definitions of embodiment and have not focused on mental simulation as the basis for creativity (see Frith, Miller, & Loprinzi, 2019 for a review). For instance, Friedman and Förster (2000, 2002) showed that arm flexion, compared to arm extension, enhanced insight problem solving, facilitated DT generation, and improved retrieval of verbal solutions from memory on a letter string completion task. The authors interpreted their findings as supporting the hypothesis that positive state associated processing cues, such as arm flexion, facilitate creative performance by diminishing retrieval blocking (Friedman & Förster, 2002). However, this result also shows that the state of the body (the flexion of the arms in particular), influences DT. A reasonable hypothesis is that this occurs to do motor simulation priming. In another type of investigation, Leung et al. (2012) examined how metaphors of embodiment can facilitate creative performance. The authors showed that gesturing with each hand and putting objects together enhances performance on DT. However, while these studies demonstrate the linkage between creative performance and bodily states, they do little to illuminate the strategies underlying creative output.

It is important to note that findings from neuroimaging, while not definitive evidence of embodied or grounded strategies, provide additional insights into potential mechanisms if we allow for reverse inference. First, previous research has shown that generating uncommon uses is correlated with activations of posterior regions implicated in the visual analysis of objects (Chrysikou & Thompson-Schill, 2011). Further, patterns of activation within dorsal regions implicated in organizing and selecting actions towards objects reflect action information more during the generation of uncommon uses (Matheson, Buxbaum, & Thompson-Schill, 2017). Benedek et al. (2014) showed that the inferior parietal lobe, including a part of the supra marginal gyrus, a region implicated in tool use, was more active during the generation of novel creative uses. Aziz-Zadeh et al. (2013) showed that a network of premotor and parietal areas that are involved in a visuospatial creativity task; given the role of these regions in preparing motor actions, the authors suggest that the participants simulate possible futures and use motor imagery to complete the task. A recent meta-analysis showed that posterior 'perceptual' regions including the fusiform gyrus and the parietal lobe are consistently active in visuospatial creativity tasks (Pidgeon et al., 2016; Wu et al., 2015). Finally, both functional connectivity (Cousijn, Koolschijn, Zanolie, Kleibeuker, & Crone, 2014; Cousijn, Zanolie, Munsters, Kleibeuker, & Crone, 2014) and anatomical connectivity (Kenett et al., 2018) studies have found evidence of activation in motor regions during DT. Taken together, one hypothesis is that posterior cortical regions, given their correlated activity in basic visual and motor tasks (i.e. sensorimotor tasks), support creativity by implementing visual and action simulations (see also Fink et al., 2010). These simulations are exactly those proposed by the embodied cognitive framework.

In summary, previous research suggests that the embodied framework may offer insights into the cognitive strategies people use in DT. In the present exploratory study, we developed

a novel coding scheme that attempts to capture aspects of the cognitive strategies that participants use to generate uncommon uses. This was achieved by characterizing participant DT responses on a number of dimensions derived from considerations of the behavioural, neural, and theoretical background reviewed here, as well as our own observations of responses in this task collected in a previous study (Matheson et al., 2017)[1]. To develop these dimensions, we first assured their reliability in a sample of independent raters. Then we explored, using previously published DT responses (Silvia et al., 2017), how the use of these dimensions changes as the number of responses generated by participants increases. Typically, as the number of ideas increases, the originality of ideas increases, an effect thought to reflect executive changes that occur as the task unfolds (i.e. the serial order effect; see Beaty & Silvia, 2012), specifically task switching or inhibition (Acar & Runco, 2017; Wang, Di, & Qian, 2007). Despite these insights, the details of the additional mechanisms underlying this effect remain unknown. Additionally, we investigated how the use of these different dimensions varies as a function of individual differences in creativity. To do so, we used participants' scores on the Creativity Achievement Questionnaire (CAQ), a subjective questionnaire measuring creative output in a number of domains (Carson, Peterson, & Higgins, 2005). Using these dimensions, we provide a preliminary characterization of potential strategies people use when generating uncommon object uses, how these strategies change over time, and how they relate to individual differences in creative ability. Specifically, we explore the extent to which responses in this task are characterized by the use of these dimensions, how different strategies arise as the demand to be creative increases, and in what way creative people use these strategies.

Methods

Materials

We developed our dimensions with four complimentary approaches: a) We interrogated the literature on embodied tool use and embodied conceptualization generally, b) we reviewed the existing literature on cognitive strategies in DT, c) we reviewed the neuroimaging literature on visual and motor imagery and grounded cognition, and d) we qualitatively analyzed DT responses from a previous study (Matheson et al., 2017). Together, these considerations allowed us to develop a list of potentially useful dimensions related to DT. Note that some of our dimensions were inspired more by *a priori* approaches (e.g. egocentric vs. allocentric, toward vs. away dimensions), while others were more inspired *a posteriori* in a data driven fashion (e.g. the adverb vs. conjunction dimension). These dimensions were developed by identifying dichotomous, mutually exclusive features of participant responses. While certainly none of these dimensions are conclusively or exclusively 'embodied', they reflect variables that are especially relevant to the embodiment framework (Matheson & Barsalou, 2018). For instance, they reflect the presence or absence of modality related information, they include the body as a foundation or outcome of the given creative response, or they appear to reflect aspects of action or visuomotor simulation. Thus, for the

---

[1]Of course, there is no 1:1 mapping of the details contained within verbal reports and the underlying cognitive mechanisms that support verbalizing in this task (see Khandwalla, 1993). However, as discussed in the discussion, the dimensions we characterize here at least imply particular cognitive processes over others.

most part, these dimensions are quite unlike previously identified dimensions (e.g. novelty) and may offer a unique way of characterizing divergent thinking responses.

The initial dimensions and how they relate to existing theoretical ideas or empirical findings are described below.

The dimensions

1.   Analogy vs. action.—Much has been written about the use of embodied metaphors for understanding abstract concepts. That is, we often map concrete, embodied experience (e.g. walking down a path) to another domain (e.g. human relationships) to understand that domain (Lakoff & Johnson, 2008). Thus, it is possible that we might creatively use one object in analogy or as a metaphor for another. In scrutinizing participants' responses collected by Matheson et al. (2017), responses were often given in such a way that the action is described as an analogy or metaphor with an adverb/conjunction (i.e. 'as' or 'like'); for example, "Use a brick as a doorstop," or "Use a brick like an instrument." Here, experiences in one domain (i.e. experiences of brick-ness) are mapped to a new use (i.e. experiences as doorstop-ness). Conversely, some responses are given with an explicit verb; for example, "Use a brick to smash a window." Here, a simulation of a specific action might underlie the response, and no aspect of experience is mapped to a new domain; that is, the novel use is understood as its own concrete experience. This may reflect the use of a simulation of a specific action that is not mapped to the new domain. Overall, people may generate creative uses by either simulating analogous uses (of one object mapped to another) or by simulating explicit, specific actions.

2.   Whole vs. object part.—Previous research has established that people will often verbally decompose the object into parts (Gilhooly et al., 2007) while generating creative uses. Decomposing the object requires a careful analysis of the visual and material form of the objects and the actions these parts afford. In general, thinking about the properties of an object (e.g. shape vs. colour) activates corresponding modality-related cortices (Oliver & Thompson-Schill, 2003), a finding consistent with the idea that simulations within modality-related cortices support thinking about objects. Indeed, in scrutinizing DT responses collected by Matheson et al. (2017), some responses were given in such a way that the action describes using decomposed parts of the object; for instance, "Use the leg of the chair to dig in dirt." Conversely, other responses describe the entire object; for example, "Use a chair to smash a window." This suggests that people may generate creative uses by simulating affordances of either parts of objects or more wholistic forms of them.

3.   Same vs. different action.—Much research has investigated the putative 'tool-use' brain network (Gallivan, McLean, Valyear, & Culham, 2013; Johnson-Frey, 2004; Lewis, 2006). Within this network, brain regions such as the supramarginal gyrus have been implicated in specifying different (often conflicting) action possibilities towards objects; it is thought that the resultant action is selected from these possibilities (Watson & Buxbaum, 2015). In scrutinizing the DT responses of participants collected by Matheson et al. (2017), some responses were given in such a way that the action described is the same as the action that would be used for the object's common use. For instance, a hammer is commonly used

with a hammering action, and so "Use a hammer to smash a window" describes the same action, though with a different possible outcome. Alternatively, some responses are given such that a different action is used. For instance, "Use a hammer to roll dough" describes an action that is different from the common use. This suggests that people may generate creative uses by simulating multiple actions that the object affords, some of which are the common actions while some are additional possible actions. The creative use then is based on a common or different action.

4.   Egocentric vs. allocentric.—Previous behavioural research has demonstrated that different action representations are activated when viewing objects from an egocentric compared to an allocentric perspective (Bruzzo, Borghi, & Ghirlanda, 2008). This is also supported by neural imaging results showing increased activity in sensory-motor cortices when participants viewed egocentric actions compared to more occipital regions during allocentric actions (Jackson, Meltzoff, & Decety, 2006). Consistent with this dichotomy, participants' DT responses collected by Matheson et al. (2017) were given such that some describe an action form the first person, egocentric perspective; for instance, "Use a hammer to support my weight like a cane". Others are written from a third person, allocentric perspective; for instance, "Use a hammer to tenderize meat". This suggest that people may simulate creative uses from either an egocentric 1st person perspective (i.e. as though they were performing the action and describing the outcome), or from a 3nd person perspective (i.e. as though they were watching someone perform the action).

5.   Concrete vs. abstract.—It is well established that reading concrete nouns activates dissociable neural networks (Binder, Westbury, McKiernan, Possing, & Medler, 2005) and reading action verbs (e.g. kick, pick, lick) activates somatosensory cortex in a topographic manner (Pulvermüller, Härle, & Hummel, 2001), a finding that is often argued to reflect a simulation of bodily action to understand the verb. Further, words that tend to refer to things that are experienced with the body specifically activate sensory-motor cortex (e.g. the supramarginal gyrus; Hargreaves et al., 2012). The DT responses people give clearly reveal variance in the concreteness of their verbs or nouns. For instance, in the data collected by Matheson et al. (2017), some responses are given in such a way that the action and objects are concrete, clearly specified, or easy to (visually or motorically) imagine; for instance, "Use a hammer to hammer a nail." Here, the verb 'to hammer' and the noun 'nail' are concretely specified. Others are given in such a way that the action is more abstract, under-specified, and difficult to imagine, or there may be multiple actions that support the goal; for instance, "Use a hammer to make art". Here the verb 'to make' is underspecified (there are potentially infinite actions a person could take in the process of 'making') and the object 'art' is abstract (art could be a painting, a song, a poem, or potentially infinite things). This suggests that people may generate creative uses by simulating a very specific, concreate action, or do so only vaguely with an underspecified action. (Note that this dimension could be labeled *Specific vs. unspecific*, but we will use our labeling for simplicity. It overlaps conceptually with the *Analogy vs. action* dimension, in that concreteness defines one pole of the dimension).

6.   Novel vs familiar.—Previous research on DT has shown that, in their initial attempts to generate a creative use, participants tend to retrieve familiar uncommon uses (e.g. use lipstick as a writing tool; Silvia et al., 2017) which are familiar to them (and others) and do not reflect true, uncommon uses that are generated creatively 'online'. Neuropsychologically, it is known that retrieving information from long term memory dissociates from generating actions 'online' (i.e. the two-streams hypothesis: Milner & Goodale, 2008). This suggests people may simulate properties of actions and objects that are based on long-term representations or may rely more on online simulations pertinent to information available to them in real time. Because this dimension has been observed elsewhere (Silvia et al., 2017), we have included it here.

7–9.   Three types of action posture.—With respect to common uses of objects, previous research has established that three dimensions are important for characterizing the types of postures different objects afford (Watson & Buxbaum, 2014). Specifically, this research characterizes a) how much surface area of the hand interacts with the object, b) how much arm movement is involved in producing the action, and c) whether the action is towards or away from the body. Given that people may simulate possible actions with the objects in generating creative uses, we intuited that these characteristics and may be relevant. Specifically, participants may generate creative uses by simulating these highly specific features of actions, such as their relationship to the rest of the body or the muscle contractions involved in prehension.

The coding scheme and data for coding

From these dimensions, we created a coding scheme (see Appendix) that naïve raters could use to rate DT responses. These responses were taken from a study by Silvia, Nusbaum, and Beaty, (2017). In that previous study, 151 participants continuously generated as many alternative uses of a 'box' they could in three minutes. Furthermore, each participant in that study completed the creative achievement questionnaire to assess individual differences in creative ability (Silvia et al., 2017). The Creative Achievement Questionnaire (CAQ; Carson et al., 2005) measures real-world creative accomplishments in 10 domains: visual arts, music, dance, architectural design, creative writing, humor, inventions, scientific discovery, theater/film, and culinary arts. Each domain is measured with 7 items. The first item for each domain is a "no creativity" response: participants can indicate that they have no accomplishment in the area. The items then increase in steps toward greater accomplishment, with different score weighting for some of the items. Participant scores for each domain were recorded by adding all responses for that category based on the weight of each question. The scores can range from 0 (*having no training or talent in the area*) to 28 (*selecting all categories all the way up to "My work has been recognized nationally"*). Most participants receive low CAQ scores, resulting in a skewed distribution of sample results (Silvia, Wigert, Reiter-Palmon, & Kaufman, 2012). As such, a log-transform is applied on the sum of all CAQ sub-domains to generate a general CAQ score.

Study Layout

Our investigation had two phases. First, we investigated the reliability of ratings (reliability phase) and then we investigated the use of the reliable dimensions in the data set (investigation of use phase). We describe each below.

Reliability procedure and analytical strategy

Two independent raters were given the coding scheme, and a spreadsheet containing the DT responses generated by the 151 participants reported in Silva et al. (2017). In the study of Silva et al. (2017), participants generated continuous DT responses for three minutes. For our reliability procedure, we had raters rate the first DT responses only (i.e. 151 DT responses were rated in total, one from each participant in the original study). In a spreadsheet, columns were set up with headers corresponding to each dimension[2]. Our raters were instructed to rate, on a binomial scale, each DT response on each dimension. Ratings took between 1.5 and 3 hours over two sessions.

We wanted to ensure that our coding scheme was effective and could be reliably applied by naïve and independent raters. We reasoned that, if the coding scheme was clear and effective, the independent raters should show high agreement on their ratings (e.g. there should be agreement about whether a response was concrete, for instance). To determine whether the dimensions were clear enough that they could be independently rated, Cohen's κ was calculated for each of the dimensions using the *fmsb()* package in R (*v.* 3.3.3; www.r-progect.org). We also examined the proportion of responses and the standard deviation as a way of exploring the observed tendency of our sample to use one anchor of the dimension or another. The standard deviation gives an indication of the variance observed across all the responses rated.

Investigation of use procedure and analytical strategy

The reliability procedure determined that some dimensions could not be reliably rated; therefore, we removed these from the final coding scheme[3]. We then used the remaining dimensions to conduct an exploratory investigation. We aimed to explore three questions. Specifically, we explored the extent to which DT responses in this task are characterized by the use of these dimensions, how different strategies arise as the demand to be creative increases, and in what way creative people use these dimensions. For instance, it may be the case that upon encountering the instruction to generate as many creative uses as possible of a shoe, participants tend to adopt the use of concrete simulations of action initially but eventually resort to the use of analogies (or vice versa), and perhaps this is especially pronounced in creative individuals. Again, this analysis was exploratory, and we had no strong *a priori* hypotheses regarding how the use of these dimensions would change, only that they may. However, we reasoned that such changes in the use of each dimension would

---

[2]For thoroughness, we also initially included four additional dimensions identified by Gilhooly et al. (2007) though we do not discuss them further.

[3]Note that these included the property dimension of Gilhooly et al. (2007), where the ratings by our two independent raters were almost perfectly opposite. Further, there was very little variance on other dimensions (e.g. memory use, where none were deemed to be based on memory). While we intended to compare these dimensions with the ones we collected, the lack of agreement and variance made this difficult. Because replicating findings based on these dimensions was outside of the scope of our investigation, they will not be discussed further.

provide preliminary evidence that these dimensions are useful in characterizing the strategies participants use in DT and how they change across time in relation to individual differences in creative ability.

To that end, we analyzed the first five DT responses generated to the word 'box' for all 151 participants collected by Silvia, Nusbaum, and Beaty (2017)[4]. In this dataset, 100% of the participants generated one response, 97% generated two responses, 95% generated three responses, 88% generated four responses, and 80% generated five responses. To obtain independent ratings of these responses, we split them into 10 sub-lists by a median split for each of the DT responses (e.g., two lists for DT responses generated as the first response, each including 50% of the responses, two lists for DT responses generated as the second response, etc.). Identical responses generated by different participants were merged so that each list contained only unique responses. Each one of these sub-lists were submitted, via qualtrics (www.qualtrics.com), to an Amazon Mechanical Turk rating survey (AMT; Buhrmester, Kwang, & Gosling, 2011). This is motivated by a recent study showing how naïve AMT raters can reliably rate divergent thinking responses (Hass, Rivera, & Silvia, 2018). Each survey was sent to 50 AMT raters who rated all DT responses in their list; they rated each response on the remaining seven dimensions. All raters rated only one sub-list. Each AMT rater downloaded a CSV file that included the DT responses in one column, headers for each of the seven dimensions, and the detailed ratings instructions (Appendix). AMT raters were encouraged to complete the ratings with the instructions at hand, and to quickly, without over-deliberating, whether they thought the DT response was a 1 or a 2 (depending on the dimension). AMT raters were also instructed to indicate 3 if unsure.

The AMT rating data were re-coded as a 1 or a 0. 3 (i.e. unsure) ratings were discarded from analysis. Two statistical procedures were used to analyze the ratings. First, for descriptive purposes, proportions of ratings (e.g. analogy responses vs. verb) were submitted to a one-way between-participants ANOVA with DT response order as a factor. With this analysis, we aimed to provide a simple test of whether the overall proportion of ratings favoring one dimension changed over time. We visualized this analysis by plotting the proportion of rated DT responses as a function of DT response order.

We additionally conducted a more detailed, second, analysis. In this analysis, a binomial generalized linear mixed effects model comparison procedure was used to explore changes to the probability of the rating of each DT response as a function of both DT response order and the CAQ standardized score of the participant who gave the DT response. To do so, we used the *glmer()* function of the *lme4()* package (Bates et al., 2015). Generalized mixed models have a number of advantages over the traditional ANOVA when dealing with categorical data, including using all of the data (rather than aggregating) to find the optimal model; they also allow us to not violate the assumption of the homogeneity of variance, which is almost certainly violated when dealing with proportions (Jaeger, 2008). Further, mixed models allow us to treat all variables as continuous, including the responses order (whereas the ANOVA forces us to treat response order as a categorical variable). This is

---

[4]Of course, participants can generate more than 5 alternative or creative uses of a box. However, in the dataset we used here, the percentage of participants generating responses drops drastically after 5, and very few reach higher numbers (see also Hass, 2016).

important because the values of the DT response order are meaningful; it is more difficult to produce a DT response after having just given a previous DT response; additionally, the semantic content of a given DT response changes predictably from closely related content to more semantically distant content (Hass, 2017b). Thus, while response order is technically an ordinal variable, there are pragmatic advantages to treating it as a continuous variable that captures changes in cognitive strategies from time point to time point. Overall, our analysis strategy better captures the role of the predictors.

For each dimension, we fitted a base model with DT response order as a fixed effect (continuous predictor) and AMT rater as a random intercept (to account for each rater's individual bias or perhaps other variability due to individual differences). Note, to improve interpretability of the intercept, we coded response as 0–4 rather than 1–5 in this analysis. We compared this fit to another model with the same random effect structure but with an interaction effect between DT response order and CAQ. To improve the interpretability of the coefficient values for CAQ in the reporting of the results, we standardized CAQ values for this analysis; thus, the resulting value reflects the coefficient for the average CAQ score. This allowed us to determine whether adding the CAQ score improved the fit of the model.

We visualized this analysis by plotting the estimated slopes at four different values of the *unstandardized* log-transformed CAQ score (i.e. not the z scored values). We visualized slope estimates at the 1st quartile (log-transformed CAQ = .5), median (log-transformed CAQ = 1), 3rd quartile (CAQ = 1.2) and maximum (log-transformed CAQ = 1.8) allowing for an intuitive legend capturing the varying levels of CAQ. These slopes estimate the change in the likelihood of the rating (e.g. the likelihood a DT response was an analogy) as a function of response and CAQ. These were extracted using the *effects()* package (Fox, 2003).

We report both the ANOVA and the results of the model comparison procedure using the *anova()* function in R. Further, we used the methods of Nakagawa and Schielzeth (2013) to provide estimates of the marginal R^2 (i.e. variance explained by the fixed effects) for all models using the *MuMIn* package (*v.* 1.40).

Results

Our first objective was to determine whether the instructions in our coding scheme and its dimensions were clear enough that they could be used to rate DT responses in the alternative uses task reliably by different raters.

Reliability

Cohen's κ was computed for each dimension along with the proportion of ratings and the standard deviation (as an indication of variance) across items (Table 1). We found almost fair to perfect agreement on most dimensions with the exception of *Egocentric vs. allocentric* and *Arm movement*. For *Egocentric vs. allocentric*, the raters did not vary on their ratings (no variance) and had perfect agreement. Both of our raters claimed that every DT response was from an allocentric perspective. It is unclear whether this is because all DT responses are truly described in an allocentric nature, or all DT responses failed to include enough

detail for this dimension to be useful. Also, it is unclear how perspective might be expected to change over the duration of the task or how readily information about perspective could be expected from verbal responses. For *Arm movement*, one rater claimed all DT responses involved substantial movement of the arm, while the other claimed almost the opposite. This may be due to a lack of sufficient detail in the DT responses themselves that were rated. While future research may provide more fruitful outcomes for these dimensions, especially for cues other than 'box', these dimensions were excluded from further analysis.

Exploratory investigation

Some AMT raters failed to give ratings that confirmed to our instructions; this occurred as some raters accidentally and occasionally failed to follow our specific instructions of coding as 1 or 2 (e.g. gave T or F, 'yes' or 'no', or sometimes included values that were not requested). After removal of these AMT raters and data, the resulting N for each response (first, second, etc.) was as follows: 1 = 83; 2 = 66; 3 = 74; 4 = 59; 5 = 56. Proportions (M) and standard deviations (SD) of ratings on each dimension for each of the first five DT responses are provided in Table 2. Due to some responses being coded by AMT raters incorrectly for different dimensions (with values other than 1, 2, or 3), or that for some dimensions some participants in the original study did not generate five DT responses, the ANOVAs for each dimension has slightly different degrees of freedom. The complete interaction models for each dimension are presented in table 3. Note also, that thought the marginal R^2s were small in these analyses, the conditional R^2s (variance explained by both fixed and random effects) were all around .30.

Analogy vs. action.—The ANOVA revealed a non-significant effect of response order, $F(4,333) = 1.4$, $p = .23$, $\eta_g^2 = .02$ (Figure 1). The model comparison procedure revealed that a model with the Response order X CAQ interaction model was a better fit ($DF = 5$, BIC = 21098, $R^2_m = .004$) than the model with only response order ($DF = 3$, BIC = 21115, $R^2_m = .001$), $X^2(2) = 36.9$, $p < .001$. This model revealed a significant interaction between response order and CAQ, $b = .06$, SE = .01, $z = 4.86$, $p < .001$. This interaction indicates that the change in the use of analogies depends on creativity (Figure 2). High creative individuals were more likely to increase their use of analogies as they produce more DT responses. Conversely, low creative individuals are more likely to reduce their use of analogies as they produce more DT responses.

Whole vs. parts.—The ANOVA revealed a significant effect of response order, $F(4,334) = 4.89$, $p < .001$, $\eta_g^2 = .06$ (Figure 1). However, the GLME models failed to converge and could not be fit. This is likely due to the uneven distribution of whole vs. parts responses. In the present data, 16% of the responses consisted of describing uses that specified an object part. Adjustments to GLME optimizers or to the link function would not result in a fit; therefore, no effects were calculated for this data.

Same vs. different action.—The ANOVA revealed a non-significant effect of response order, $F(4,334) = 1.48$, $p = .21$, $\eta_g^2 = .02$ (Figure 1). The model comparison procedure revealed that a model with the Response order X CAQ interaction model was a better fit ($DF = 5$, BIC = 20885, $R^2_m = .003$) than the model with only response order ($DF=3$,

BIC = 20895, R^2_m = .001), X^2 (2) = 30.23, $p < .001$. This model revealed a significant interaction between response order and CAQ, $b = -.07$, SE = .01, $z = -5.3$, $p < .001$. This result suggests that the use of same responses depends on creativity (Figure 2). Low creative individuals tend to show less change and are more likely to rely on same responses throughout the response period; conversely, high creative individuals tend to produce more responses that describe same actions initially but then rely on different responses to a greater extent later on.

Concrete vs. abstract.—The ANOVA revealed a significant effect of response order, F(4,334) = 2.66, $p = .03$, $\eta_g^2 = .03$ (Figure 1). The model comparison procedure revealed that a model with the Response order X CAQ interaction model was a better fit ($DF = 5$, BIC = 23018, R^2_m = .004) than the model with only response order ($DF = 3$, BIC = 23029, R^2_m = .002), X^2 (2) = 30.91, $p < .001$. This model revealed a significant interaction between response order and CAQ, $b = -.05$, SE = .01, $z = -4.3$, $p < .001$. This result suggests that the tendency to produce concrete responses depends on creativity (Figure 2). While both low and high creative individuals are more likely to decrease their use of concrete responses, high creative individuals show a stronger tendency to do so.

Novel vs. familiar.—The ANOVA revealed a non-significant effect of response order, F(4,334) = .85, $p = .49$, $\eta_g^2 = .01$ (Figure 1). The model comparison procedure revealed that a model with the Response order X CAQ interaction model was a better fit ($DF = 5$, BIC = 24567, R^2_m = .002) than the model with only response order ($DF = 3$, BIC = 24576, R^2_m = .0001), X^2 (2) = 28.2, $p < .0001$. There was a marginally significant interaction between response order and CAQ, $b = -.02$, SE = .01, $z = -1.68$, $p = .09$.

Towards vs. away.—The ANOVA revealed a non-significant effect of response order, F(4,329) = .38, $p = .82$, $\eta_g^2 = .005$ (see figure 1). The model comparison procedure revealed that a model with the Response order X CAQ interaction model was a better fit ($DF = 5$, BIC = 21449, R^2_m = .003) than the model with only response order ($DF = 3$, BIC = 21453, R^2_m = .001), X^2 (2) = 23.6, $p < .001$. The model included a significant interaction between response order and CAQ, $b = .03$, SE = .01, $z = 2.27$, $p = .02$. This result suggests that creativity modulates the likelihood a person will describe an action that is towards their own bodies (Figure 2). Specifically, in low creative individuals, the probability that they will describe an action towards themselves starts higher and decreases as they generate uses. Conversely, the probability of a towards response is lower in high creative individuals and does not change throughout the response period.

To explore potential confounds with CAQ, we performed a pairwise correlations analysis on CAQ and participant age and participant's fluency in the task (i.e. the number of responses they generated in the previous study[5] (Table 4). Note that none of the correlations were strong, failing to reach significance. Thus, our results are likely not due to differences in participant age or overall fluency.

---

[5]Note that 13 participants failed to include their age and were excluded from the correlation analysis.

We also explored the relationship between our dimensions in a post hoc exploratory analysis. To do so, we calculated the pairwise phi correlation within each dimension and tested significance with the Chi-square test. We used the *xtab_statistics()* function from the sjstats library in R (Lüdecke, 2018). The largest correlations come from the pairwise pairs of same vs. different, concrete vs. abstract, and novel vs. familiar (Table 5).

Discussion

Divergent thinking (DT) tasks are widely used to measure creativity in creativity research (Acar & Runco, 2019; Runco & Acar, 2012). However, the strategies participants use when generating, for example, alternative uses to objects is far from understood. In the present study, we devised a novel coding scheme for rating DT responses with considerations from the behavioural and neural literature on tool-use and theoretical considerations from embodied or grounded cognition. With it, we provide a preliminary characterization of verbal responses that suggest potential cognitive strategies people use when generating creative object uses. The results from our exploratory investigation reveal that responses are split with respect to the dimensions we identified and can be reliably coded by naïve raters. Further, we provide a preliminary characterization of how these strategies change over time and how they relate to individual differences in creative ability.

Our dimensions were able to characterize strategies people used to generate DT responses and they could be reliably coded by naïve raters. For instance, in response to the instruction to describe as many creative uses of a box as they could, about half the participants provided responses that described actions towards the body while the other half described actions away from the body. However, some dimensions showed larger proportions in favor one value of the dimension over the other. For instance, most participants provided responses that described the use of the whole object, and overall there was a greater proportion of responses describing actions that were different from the common use actions. Capturing the variability in using these dimensions is an important step towards understanding the strategies people use to generate their responses. Importantly, these dimensions were developed to be mutually exclusive (i.e. if a response is abstract it cannot be concrete) and therefore shed light on some of the dichotomous dimensions people use in generating their responses. Further, discovering that naïve raters could reliably code responses based on these dimensions provides a unique contribution to understanding performance in the divergent thinking task.

However, reliable coding of these dimensions does not illuminate the ways in which different individuals are predisposed to use any given strategy. To address this, we investigated whether individual differences in creativity (as measured by the CAQ), predicts how people use these dimensions over time in generating DT responses. Our results provide preliminary support that more creative people use these dimensions differently over time than less creative people. For each dimension (with the exception of whole vs. parts, where the data did not allow us to fit a model, and novel vs. familiar, where the interaction was only marginally significant), creativity interacted with response number, revealing how the use of these embodied dimensions change over time and as a function of creativity.

From these changes, we can generate a number of novel speculations and hypotheses about the cognitive strategies that people use in this task. We summarize the main pattern of results for each dimension and detail our speculations below.

Analogy vs. action

High creative individuals tend to rely more on analogies later in the response period, while low creative individuals tend to rely on them to a lesser extent. For instance, one participant with a high CAQ log-transformed score (i.e. 1.6) first gave the response 'hold books' (a response that expresses a more specific verb). However, by their 4th response, they suggested that it could be used 'as a doll house' (a clear analogy). In contrast, a participant with a low CAQ log-transformed score (i.e. 0) gave the first response as 'as a makeshift home' while their 5th response was 'as a trashbag', both that clearly indicate a persistent use of an analogy strategy later in the response period.

The increase in the use of analogies in high-creative participants suggests that a gateway towards creativity may be thinking about specific aspects of the situation at hand and mapping them to a new domain. There is evidence that an important step in creative problem solving is identifying all of the components available in any given situation (Abdulla, Paek, Cramond, & Runco, 2018; McCaffrey, 2012). For instance, in thinking about uses of a candle, one might recognize that there is a wick, which is made of string, which is really just a strand of fibourous material. In this way, creative individuals may push themseles beyond thinking of the concrete properties of objects in a typical context, and extend it to a new context. This may not be the case in low creative individuals. Analogical thinking has been related to creativity more generally (Gick & Holyoak, 1980; Green, 2016), and here we show that high creative individuals may find analogous uses of objects more accessible.

Same vs. different action

High creative individuals are more likely to generate responses that use different actions later in the response period. For instance, a participant with a high CAQ log-transformed score (i.e., 1.54) first responded with 'to hold books', a common use of a box. By the 5th response, however, they suggested that it could be used 'as a doll house', a description that requires many different actions to execute (i.e. different from simply putting things in boxes). In contrast, low creative individuals tend to produce more common responses throughout the response period. For instance, one participant with a low CAQ log-transformed score (i.e., 0) suggested using 'as a shoe' for the 1st suggestion and by the 5th response was generating 'carrying things', both more common actions associated with boxes (i.e. putting things in them).

This tendency is likely related to issues of functional fixedness, that is, the difficulty in thinking about different functions of objects (Duncan, 1945). It is widely recognized that a challenge in overcoming a problem involves reducing or eliminating functional fixedness (e.g. 'restructuring'; Ohlsson, 1992) and one way to do so might be to change the modality of the stimuli (e.g. pictures result in more functional fixedness than words; Chrysikou et al., 2016). Our results extend this, demonstrating that high creative individuals are adept at finding responses in which the actions differ from the common actions, especially later on.

This may imply that one route to overcoming functional fixedness is to focus explicitly on the actions that are typically performed with an object and deliberately attempt to generate novel, different actions.

Concrete vs. abstract

We also observed that high creative individuals use fewer concrete verbs later in their responses, a phenomenon possibly related to the fact that high creative individuals rely more on analogies later in responding. Importantly, imagery does predict creative cognition (Abraham & Bubic, 2015; LeBoutillier & Marks, 2003). In adults, inducing imagery can increase creative performance on divergent thinking tasks like the one used here (Durndell & Wetherick, 1976); in children, which have been more widely studied, inducing imagery exercises increases creative output on a wide range of tasks (Antonietti & Colombo, 2011). These results show that individuals are likely to activate concrete representations of actions or objects, but creative individuals abstract away from concrete experience later in responding. However, it is unclear whether the more abstract descriptions given by creative individuals are their most creative output.

Novel vs. familiar

Though for the novel vs. familiar dimension the interaction between CAQ and Response Order was not significant ($p = .09$), there are hints that the use of truly novel (as opposed to common) responses depended on individual differences in creative ability. This may be because this dimension overlaps largely with the same vs. different dimension (as shown in our correlations). We attempted to differentiate between actions that participants may describe (same vs. different actions; i.e. swinging vs. pulling with a hammer) and the functions that different objects serve (novel vs familiar; i.e. to put nails into the wall vs. to spear fish). It may be the case that these dimensions are not independent when describing actions in response to 'box' though they may be for other divergent thinking cues (e.g. hammer, soap dispenser, mop, etc.). Future research should assess this possibility (Hass, 2017b).

Towards vs. away

Finally, only one dimension described in the action semantics of Watson and Buxbaum (2014) was reliable enough to use as a dimension describing DT strategies, revealing an interaction between CAQ and response order. Participants varied in their use of descriptions of actions (or their goals) that were directed either towards or away from their own bodies. Our results suggest that individuals lower in creativity are more likely to describe actions or their goals towards their body in earlier responses, while higher creative individuals are more stable in their use of this dimension. For instance, a participant with a high CAQ log-transformed score (i.e., 1.32) suggested using the box 'as a hat' for the 1st response and 'a form of clothing' on their 4th response, both of which are directed at the body or have goals directed towards the body. Conversely, one participant with a low log-transformed CAQ score (i.e., 0) initially generated the response that the box could be used as a 'hiding spot' (an action that directs the box towards the body) but in later responses suggested using the box 'as a basketball rim' (an action with a goal directed out in the world).

This result is interesting for a number of reasons. Previous research has shown that appetitive motor movements, including flexing the arms towards the body, result in superior performance on DT compared to avoidant motor movements away from the body (Friedman & Förster, 2002). Previous authors have speculated that appetitive movements, after a history of positive reinforcement (e.g. through eating), induce positive affective states that have consequences for cognitive performance, including creative output (Cacioppo, Priester, & Berntson, 1993). Importantly, however, the effects of flexion and extension on creativity exist over and above any conscious experience of affect (see also Friedman & Förster, 2000), suggesting that this phenomenon is not simply caused by feelings of happiness. Our results suggest that simulations of actions towards vs. away from the body differentiate low and high creative participants. One intriguing possibility is that low creative individuals can exploit these simulations to generate creative responses, especially early in the response period. The relative stability of the use of this dimension in high creative individuals suggests it is a strategy that does not need to be relied on. We suggest that such simulations make use of the omnipresence of the body as a recipient of an object-based action, and in doing so, open up more possibilities for creative uses, a strategy that might benefit lower creative individuals in particular. Considering possible movements towards one's own body highlights affordances of objects that go otherwise unnoticed compared to when simulating actions out into the world, where the recipients of the imagined object use are ever changing, inconsistent, or not relevant to object use. Future research should explore this possibility.

The use of embodiment to generate creative uses

The dimensions we have developed here contribute to our understanding possible cognitive strategies that allow people to generate creative alternative uses for objects. Previous research has emphasized memory strategies (e.g., Chrysikou et al., 2016; Chrysikou & Weisberg, 2005) or decompositional strategies (e.g., Knoblich, Ohlsson, Haider, & Rhenius, 1999; McCaffrey, 2012) and has evaluated creativity based on subjective judgments of creative output (e.g. how novel and useful is the response). While these approaches help describe important variables in predicting creative responding, our results begin to identify specific potential mechanims that give rise to creative responses more generally. A hypothesis stemming from our preliminary results is that more creative individuals are likely to use analogies, suggesting that one mechanism of generating alternative uses is the ability to activate neural simulations of specific actions and/or action components and map them to new domains. This is supported by the general finding that imagery predicts creativity (LeBoutillier & Marks, 2003). Creative individuals may also resist the temptation to describe the same actions as those that would commonly be used, suggesting that they are more prone to simulating competing actions rather than the learned associate actions (Cisek, 2007; Jax & Buxbaum, 2010). They are also more likely to produce truly novel descriptions (by definition) and they are stable in the use of their own bodies as the base for simulating actions.

Caveats

Our results are exploratory and our interpretations are speculative based on an embodied or grounded cognition framework. We feel that our findings implicate a number of important

candidate processes that underlie the generation of DT. However, there are a number of caveats to our interpretations that we wish to discuss.

A major caveat of the current study is that we are attempting to infer cognitive processes from the content of verbal responses that participants have given (Fox, Ericsson, & Best, 2011; Unsworth, 2016). Indeed, there are a number of issues with this approach, including a) the requirement to verbalize (rather than act) may affect the processes used to produce DT responses made by participants in the original study, and b) not all participants give verbal reports with the same amount of detail (Ericsson & Simon, 1980, 1983). However, the nature of our dimensions does suggest something conclusive about the strategies that are used, and we note that the alternative uses task does not require participants to directly reflect on the processes they used to generate their responses. Importantly, if a response is given as an analogy, this suggests that the cognitive processing that leads to the response was analogical. Similarly, it seems likely that any report of an action that had the participant's body as the end state (e.g. towards response) would certainly require a simulation of actions towards the person's body. However, we are sensitive that this might not always be the case. For instance, a concrete response implies a simulation of concrete actions, though a more abstract responses does not deny the possibility that at some point a concrete action was indeed simulated. Despite these concerns, we speculate that our results provide us with candidate strategies that future research can use to explore potential cognitive mechanisms underlying performance in this task.

There are a number of additional caveats to our study. First, for convenience, we only focused on participants first five DT responses (Hass, 2016). While we were able to detect interactions within these five responses, extending our analysis to later responses generated by participants may more strongly reveal changes in the use of these dimensions. Indeed, some highly creative participants are able to generate many more responses than five. Thus, future research should investigate the patterns of strategy use in higher creative people throughout the response period. Second, we had a group of independent participants rate responses generated by a separate group of participants; future research should investigate the ratings generated by participants on their own responses. Third, we used the CAQ as a measure of creativity, but there are other measures that might better differentiate the use of different strategies. Fourth, the marginal R^2 for each of the models is relatively small, suggesting that the interactions we obtain here account for only a small amount of the variance in the data; thus, the practical implications of these results are likely low and to be determined. This is not surprising. Our dimensions are not designed to compete with other important variables that capture variance in these tasks (e.g. personality variables) but are thought to help capture variance due to the use of possible mechanisms; we assume that these mechanisms are, more or less, universal. Future research should explore the extent to which these dimensions continue to capture theoretically interesting variation in creativity performance, or whether certain conditions increase the practical importance of the effects. Fifth, we are arguing that our response dimensions are mutually exclusive (e.g. the use of an analogy vs. not), though we recognize that Likert ratings could be used for a number of our dimensions, (e.g. same vs. different action, where there may be some cases of ambiguity). Though we have established some reliability in participant's ability to rate these, agreement is not perfect, suggesting that there are indeed cases that are ambiguous to raters. This

suggests that future research should explore these dimensions using a more Likert style rating system. Finally, all of the alternative uses were generated in response to the cue 'box', limiting the generalization of our results. Recent studies have shown how divergent thinking responses vary contingent on the cue (Hass, 2017a, 2017b). We hypothesize that the use of these dimensions would be more clearly demonstrated with objects that have clear common uses (e.g. hammers). Future research should explore this possibility.

Finally, our exploratory pairwise correlation analysis suggests that there is overlap between some of our dimensions. Though we have attempted to create mutual exclusivity within dimensions, we note that there isn't exclusivity between dimensions. However, these relationships point to the way in which mental simulations, when they support creativity, are multidimensional. The multidimensional nature of embodied simulations is something future research can investigate in the pursuit of fully characterizing the strategies that are adopted in this task.

Conclusion

In summary, our preliminary and exploratory findings suggest that creative individuals produce responses that are supported by the ability to map highly specific, concrete, simulations of possible actions to a new domain; these simulated actions stand in competition to the one's typically specified by the object; and are stable in their use of the body as the end state of possible creative actions. The approach we took was bottom-up and exploratory, attempting to shed novel light on the mechanisms of generating alternative uses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Abdulla AM, Paek SH, Cramond B, & Runco MA (2018). Problem finding and creativity: A meta-analytic review. Psychology of Aesthetics, Creativity, and the Arts.

Abraham A, & Bubic A (2015). Semantic memory as the root of imagination. Frontiers in Psychology, 6, 325. [PubMed: 25852626]

Acar S, & Runco MA (2017). Latency predicts category switch in divergent thinking. Psychology of Aesthetics, Creativity, and the Arts, 11(1), 43–51.

Acar S, & Runco MA (2019). Divergent thinking: New methods, recent research, and extended theory. Psychology of Aesthetics, Creativity, and the Arts, 13(2), 153–158.

Antonietti A, & Colombo B (2011). Mental imagery as a strategy to enhance creativity in children. Imagination, Cognition and Personality, 31(1), 63–77.

Aziz-Zadeh L, Liew SL, & Dandekar F (2013). Exploring the neural correlates of visual creativity. Social Cognitive and Affective Neuroscience, 8(4), 475–480. [PubMed: 22349801]

Barsalou LW (2008). Grounded cognition. Annual Review of Psychology, 59, 617–645.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B, Grothendieck G, Eigen C, & Rcpp L (2015). Package 'lme4'. Convergence, 12(1).

Beaty RE, & Silvia PJ (2012). Why do ideas get more creative over time? An executive interpretation of the serial order effect in divergent thinking tasks. Psychology of Aesthetics, Creativity and the Arts, 6(4), 309–319.

Benedek M, Jauk E, Fink A, Koschutnig K, Reishofer G, Ebner F, & Neubauer AC (2014). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. NeuroImage, 88, 125–133. [PubMed: 24269573]

Binder JR, Westbury CF, McKiernan KA, Possing ET, & Medler DA (2005). Distinct brain systems for processing concrete and abstract concepts. Journal of Cognitive Neuroscience, 17(6), 905–917. [PubMed: 16021798]

Bruzzo A, Borghi AM, & Ghirlanda S (2008). Hand–object interaction in perspective. Neuroscience Letters, 441(1), 61–65. [PubMed: 18599215]

Buhrmester M, Kwang T, & Gosling SD (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? Perspectives on Psychological Science, 6(1), 3–5. [PubMed: 26162106]

Cacioppo JT, Priester JR, & Berntson GG (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. Journal of Personality and Social Psychology, 65(1), 5. [PubMed: 8355142]

Carson SH, Peterson JB, & Higgins DM (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. Creativity Research Journal, 17(1), 37–50.

Chrysikou EG, Motyka K, Nigro C, Yang S-I, & Thompson-Schill SL (2016). Functional fixedness in creative thinking tasks depends on stimulus modality. Psychology of Aesthetics, Creativity, and the Arts, 10(4), 425–435.

Chrysikou EG, & Thompson-Schill SL (2011). Dissociable brain states linked to common and creative object use. Human Brain Mapping, 32(4), 665–675. [PubMed: 20533561]

Chrysikou EG, & Weisberg RW (2005). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5), 1134–1145.

Cisek P (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362(1485), 1585–1599. [PubMed: 17428779]

Cousijn J, Koolschijn PCMP, Zanolie K, Kleibeuker SW, & Crone EA (2014). The relation between gray matter morphology and divergent thinking in adolescents and young adults. PLoS ONE, 9(12), e114619. [PubMed: 25514366]

Cousijn J, Zanolie K, Munsters RJM, Kleibeuker SW, & Crone EA (2014). The relation between resting state connectivity and creativity in adolescents before and after training. PLoS ONE, 9(9), e105780. [PubMed: 25188416]

Duncan K (1945). On problem-solving. Psychological Monographs, 58(5), 1.

Durndell AJ, & Wetherick NE (1976). The relation of reported imagery to cognitive performance. British Journal of Psychology, 67(4), 501–506.

Ericsson KA, & Simon HA (1980). Verbal reports as data. Psychological Review, 87(3), 215–251.

Ericsson KA, & Simon HA (1983). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Fink A, Grabner RH, Gebauer D, Reishofer G, Koschutnig K, & Ebner F (2010). Enhancing creativity by means of cognitive stimulation: Evidence from an fMRI study. Neuroimage, 52(4), 1687–1695. [PubMed: 20561898]

Forthmann B, Gerwig A, Holling H, Çelik P, Storme M, & Lubart T (2016). The be-creative effect in divergent thinking: The interplay of instruction and object frequency. Intelligence, 57, 25–32.

Fox J (2003). Effect displays in R for generalised linear models. Journal of Statistical Software, 8(15), 1–27.

Fox MC, Ericsson KA, & Best R (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. Psychological Bulletin, 137(2), 316–344. [PubMed: 21090887]

Friedman RS, & Förster J (2000). The effects of approach and avoidance motor actions on the elements of creative insight. Journal of Personality and Social Psychology, 79(4), 477–492. [PubMed: 11045734]

Friedman RS, & Förster J (2002). The influence of approach and avoidance motor actions on creative cognition. Journal of Experimental Social Psychology, 38(1), 41–55.

Frith E, Miller S, & Loprinzi PD (2019). A Review of experimental research on embodied creativity: Revisiting the mind–body connection. The Journal of Creative Behavior.

Gallivan JP, McLean DA, Valyear KF, & Culham JC (2013). Decoding the neural mechanisms of human tool use. Elife, 2, e00425. [PubMed: 23741616]

Gick ML, & Holyoak KJ (1980). Analogical problem solving. Cognitive Psychology, 12(3), 306–355.

Gilhooly K, Fioratou E, Anthony S, & Wynn V (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. British Journal of Psychology, 98(4), 611–625. [PubMed: 17535464]

Goldstein A, Revivo K, Kreitler M, & Metuki N (2010). Unilateral muscle contractions enhance creative thinking. Psychonomic Bulletin & Review, 17(6), 895–899. [PubMed: 21169586]

Green AE (2016). Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. Current Directions in Psychological Science, 25(1), 28–35.

Hargreaves IS, Leonard G, Pexman PM, Pittman D, Siakaluk PD, & Goodyear B (2012). The neural correlates of the body-object interaction effect in semantic processing. Frontiers in Human Neuroscience, 6, 22. [PubMed: 22375111]

Harrington DM, Block J, & Block JH (1983). Predicting creativity in preadolescence from divergent thinking in early childhood. Journal of Personality and Social Psychology, 45(3), 609–623.

Hass RW (2016). Conceptual expansion during divergent thinking. In Papafragou A & Grodner D & Mirman D & Trueswell JC (Eds.), Proceedings of the 38th Annual Meeting of the Cognitive Science Society (pp. 996–1001). Philadelphia, PA: Cognitive Science Society.

Hass RW (2017a). Semantic search during divergent thinking. Cognition, 166, 344–357. [PubMed: 28601001]

Hass RW (2017b). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. Memory & Cognition, 45(2), 233–244. [PubMed: 27752960]

Hass RW, Rivera M, & Silvia PJ (2018). On the dependability and feasibility of layperson ratings of divergent thinking. Frontiers in Psychology, 9(1343).

Hennessey BA, & Amabile TM (2010). Creativity. Annual Review of Psychology, 61, 569–598.

Jackson PL, Meltzoff AN, & Decety J (2006). Neural circuits involved in imitation and perspective-taking. Neuroimage, 31(1), 429–439. [PubMed: 16406257]

Jaeger TF (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language, 59(4), 434–446. [PubMed: 19884961]

Jax SA, & Buxbaum LJ (2010). Response interference between functional and structural actions linked to the same familiar object. Cognition, 115(2), 350–355. [PubMed: 20156619]

Johnson-Frey SH (2004). The neural bases of complex tool use in humans. Trends in cognitive sciences, 8(2), 71–78. [PubMed: 15588811]

Kenett YN, Medaglia JD, Beaty RE, Chen Q, Betzel RF, Thompson-Schill SL, & Qiu J (2018). Driving the brain towards creativity and intelligence: A network control theory analysis. Neuropsychologia, 118, 79–90. [PubMed: 29307585]

Khandwalla PN (1993). An exploratory investigation of divergent thinking through protocol analysis. Creativity Research Journal, 6(3), 241–259.

Knoblich G, Ohlsson S, Haider H, & Rhenius D (1999). Constraint relaxation and chunk decomposition in insight problem solving. Journal of Experimental Psychology: Learning, Memory & Cognition, 25(6), 1534–1555.

Lakoff G, & Johnson M (2008). Metaphors we live by. Chicago, IL: University of Chicago press.

Landis JR, & Koch GG (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. [PubMed: 843571]

LeBoutillier N, & Marks DF (2003). Mental imagery and creativity: A meta analytic review study. British Journal of Psychology, 94(1), 29–44. [PubMed: 12648387]

Leung A. K.-y., Kim S, Polman E, Ong LS, Qiu L, Goncalo JA, & Sanchez-Burks J (2012). Embodied metaphors and creative "acts". Psychological Science, 23(5), 502–509. [PubMed: 22477105]

Lewis JW (2006). Cortical networks related to human use of tools. The Neuroscientist, 12(3), 211–231. [PubMed: 16684967]

Lüdecke D (2018). sjstats: Statistical functions for regression models. R package version 0.14, 3.

Matheson HE, & Barsalou LW (2018). Embodiment and grounding in cognitive neuroscience. In Thompson-Schill SL (Ed.), The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience (Vol. 4, pp. 357–383). Hoboken, NJ: Wiley.

Matheson HE, Buxbaum LJ, & Thompson-Schill SL (2017). Differential tuning of ventral and dorsal streams during the generation of common and. Journal of Cognitive Neuroscience, 29(11), 1791–1802. [PubMed: 28654359]

McCaffrey T (2012). Innovation relies on the obscure: A key to overcoming the classic problem of functional fixedness. Psychological Science, 23(3), 215–218. [PubMed: 22318998]

Milner AD, & Goodale MA (2008). Two visual systems re-viewed. Neuropsychologia, 46(3), 774–785. [PubMed: 18037456]

Nusbaum EC, & Silvia PJ (2011). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. Intelligence, 39(1), 36–45.

Ohlsson S (1992). Information-processing explanations of insight and related phenomena. Advances in the Psychology of Thinking, 1, 1–44.

Oliver RT, & Thompson-Schill SL (2003). Dorsal stream activation during retrieval of object size and shape. Cognitive, Affective, & Behavioral Neuroscience, 3(4), 309–322.

Oppezzo M, & Schwartz DL (2014). Give your ideas some legs: The positive effect of walking on creative thinking. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(4), 1142–1152.

Pearson J, Naselaris T, Holmes EA, & Kosslyn SM (2015). Mental imagery: Functional mechanisms and clinical applications. Trends in Cognitive Sciences, 19(10), 590–602. [PubMed: 26412097]

Pidgeon LM, Grealy M, Duffy AHB, Hay L, McTeague C, Vuletic T, Coyle D, & Gilbert SJ (2016). Functional neuroimaging of visual creativity: A systematic review and meta-analysis. Brain and Behavior, 6(10), e00540. [PubMed: 27781148]

Pulvermüller F, Härle M, & Hummel F (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. Brain and language, 78(2), 143–168. [PubMed: 11500067]

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

Runco MA, & Jaeger GJ (2012). The standard definition of creativity. Creativity Research Journal, 24(1), 92–96.

Silvia PJ, Nusbaum EC, & Beaty RE (2017). Old or new? Evaluating the old/new scoring method for divergent thinking tasks. The Journal of Creative Behavior, 51(3), 216–224.

Silvia PJ, Wigert B, Reiter-Palmon R, & Kaufman JC (2012). Assessing creativity with self-report scales: A review and empirical evaluation. Psychology of Aesthetics, Creativity, and the Arts, 6(1), 19–34.

Simonton DK (2016). Defining creativity: Don't we also need to define what Is not creative? The Journal of Creative Behavior, 52(1), 80–90.

Sowden PT, Pringle A, & Gabora L (2014). The shifting sands of creative thinking: Connections to dual-process theory. Thinking & Reasoning, 21(1), 40–60.

Torrance EP (1966). Torrance test of creative thinking. Bensenville, IL: Scholastic Testing Service, Inc.

Unsworth N (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(1), 50–61.

Vartanian O, Beatty EL, Smith I, Forbes S, Rice E, & Crocker J (2019). Measurement matters: the relationship between methods of scoring the Alternate Uses Task and brain activation. Current Opinion in Behavioral Sciences, 27, 109–115.

Vohs KD, Redden JP, & Rahinel R (2013). Physical order produces healthy choices, generosity, and conventionality, whereas disorder produces creativity. Psychological Science, 24(9), 1860–1867. [PubMed: 23907542]

Wang D, Di M, & Qian M (2007). A report on the third revision of combined raven's test (CRT-C3) for children in China. Chinese Journal of Clinical Psychology, 15(6), 559–568.

Watson CE, & Buxbaum LJ (2014). Uncovering the architecture of action semantics. Journal of Experimental Psychology: Human Perception and Performance, 40(5), 1832. [PubMed: 25045905]

Watson CE, & Buxbaum LJ (2015). A distributed network critical for selecting among tool-directed actions. Cortex, 65, 65–82. [PubMed: 25681649]

Wilken A, Forthmann B, & Holling H (2019). Instructions moderate the relationship between creative performance in figural divergent thinking and reasoning capacity. The Journal of Creative Behavior.

Wu X, Yang W, Tong D, Sun J, Chen Q, Wei D, Zhang Q, Zhang M, & Qiu J (2015). A meta-analysis of neuroimaging studies on divergent thinking using activation likelihood estimation. Human Brain Mapping, 36(7), 2703–2718. [PubMed: 25891081]
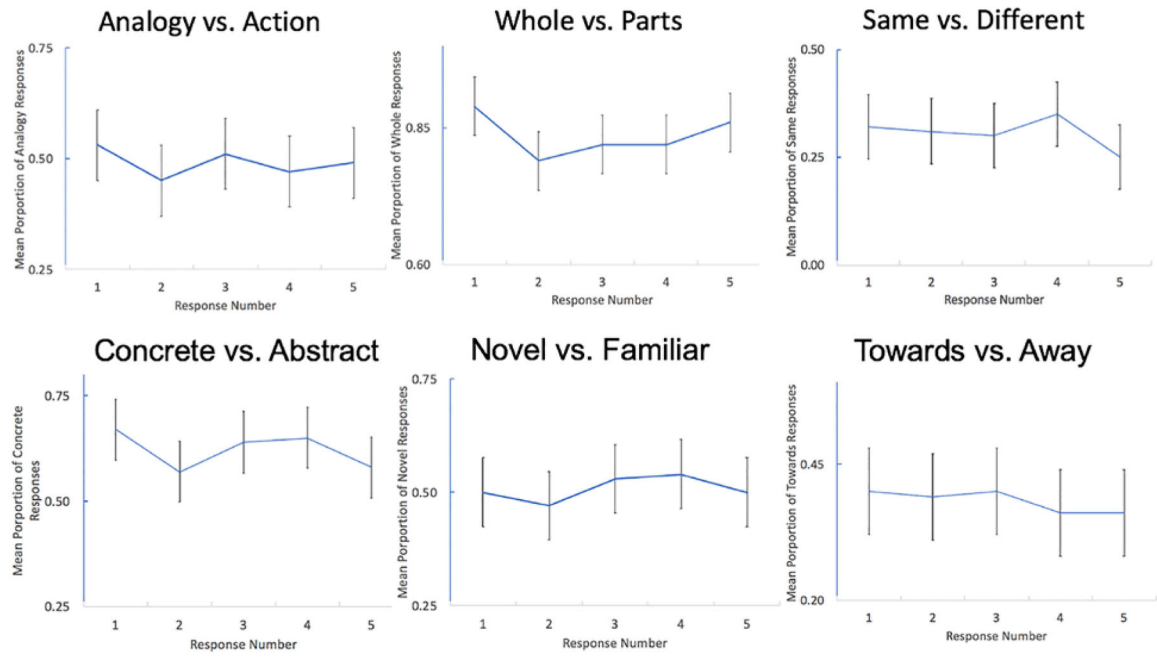
Figure 1.
Proportions for each rating as a function of response number for the different dimensions.
Error bars represent Fisher's Least Significant Difference for each pairwise comparison.
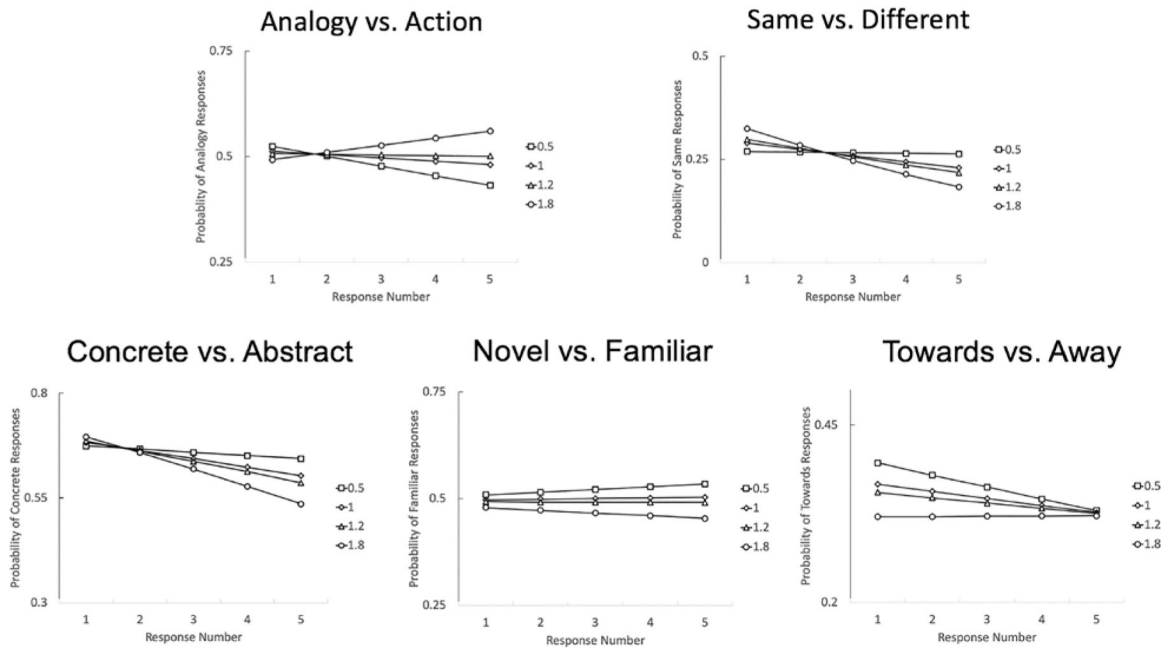
Figure 2.
Lattice figure showing the probability of each rating for the different dimensions (showing only proportions of Analogy, Same, Concrete, Novel, and Towards ratings) as a function of response order and unstandardized CAQ quartiles.

Table 1.

Cohen's κ, confidence intervals, and *p* values for inter rater reliability (2 trained raters) on each dimension. Conclusion provided by guidelines in Landis and Koch (1977).

Dimension	Proportion agreement	Cohen's κ	95% confidence interval	Conclusion
Analogy vs. action	.93	.84	.74 to .94	Almost perfect agreement
Whole vs. object part	.94	.44	.085 to .79	Moderate agreement
Same vs. different action	.97	.80	.62 to .97	Substantial agreement
Egocentric vs. allocentric	1	1	1 to 1	NA
Concrete vs. abstract	.38	.08	−.04 to .19	Slight agreement
Novel vs. familiar	.77	.42	.25 to .59	Moderate agreement
Arm movement	.15	NA	NA	NA
Hand area	.12	−.005	−.06 to .05	No agreement
Towards or away	.63	.25	.095 to .41	Fair agreement

Note - NA = not applicable

Table 2.

Proportions (M) and standard deviations (SD) for each dimension as a function of DT response order (standard deviation in parentheses). Note that proportions are calculated for the value highlighted in bold.

DT Response order	Analogy (vs. Action)	Whole (vs. Parts)	Same (vs. Different)	Concrete (vs. Abstract)	Novel (vs. Familiar)	Towards (vs. Away)
1	.53 (.23)	.89 (.15)	.32 (.26)	.67 (.21)	.50 (.23)	.40 (.24)
2	.45 (.23)	.79 (.19)	.31 (.18)	.57 (.24)	.47 (.20)	.39 (.18)
3	.51 (.26)	.82 (1.15)	.30 (.21)	.64 (.21)	.53 (.24)	.40 (.27)
4	.47 (.23)	.82 (.16)	.35 (.25)	.65 (.21)	.54 (.22)	.36 (.23)
5	.49 (.24)	.86 (.15)	.25 (.19)	.58 (.23)	.50 (.20)	.36 (.24)

Table 3.

Model estimates from the linear mixed effects models analysis for each dimension. Note, response order was coded as 0–4 allowing for better interpretability of intercept. Additionally, CAQ was standardized for this analysis.

Dimension								
	Fixed effects					**Random effect**		
		Estimate	**SD**	**z**	**p**		**variance**	**SD**
Analogy vs. action	Intercept	.07	.1	.63	.53	Subject	1.3	1.2
	Response	−.06	.05	−1.2	.23			
	CAQ	−.05	.03	−1.7	.1			
	Response X CAQ	.06	.01	4.9	<.001			
Whole vs parts	Convergence failure							
Same vs. different action	Intercept	−.94	.11	−8.4	<.001	Subject	1.4	1.2
	Response	−.05	.05	−1.0	.31			
	CAQ	.1	.03	3.4	<.001			
	Response X CAQ	−.07	.01	−5.3	<.001			
Concrete vs. abstract	Intercept	.75	.10	7.2	<.001	Subject	1.3	1.1
	Response	−.07	.05	−1.4	.15			
	CAQ	.04	.03	1.4	.17			
	Response X CAQ	−.05	.01	−4.3	<.001			
Novel vs. familiar	Intercept	.01	.10	.08	.94	Subject	1.2	1.1
	Response	.01	.04	.32	.75			
	CAQ	−.04	.03	−1.7	.09			
	Response X CAQ	−.02	.01	−1.7	.09			
Toward vs. away	Intercept	−.49	.11	−4.5	<.001	Subject	1.4	1.2
	Response	−.06	.05	−1.2	.24			
	CAQ	−.12	.03	−4.4	<.001			
	Response X CAQ	.03	.01	2.3	.02			

Note - SD = standard deviation; *z* = z score; *p* = p value corresponding to estimate; CAQ = creativity achievement questionnaire.

Table 4.

Pairwise correlations (and p values) of age, fluency, and CAQ from the data provided by Silva et al. (2017).

	Age	Fluency	CAQ
Age	1	−.08 (.*34*)	.08 (.*34*)
Fluency		1	.02 (.*82*)
CAQ			1

Table 5.

Pairwise Phi correlations (and *p* values from Chi-square test) between dimensions. Note that Phi was used given ratings are binary. Asterix indicates *p* < .05.

	Analogy vs. action	Whole vs. parts	Same vs. different	Concrete vs. abstract	Novel vs. familiar	Toward vs. away
Analogy vs. action	1.00	.06*(<.001)**	*.002(.81)*	*.02(.006)**	*.02(.003)**	*.07(<.001)**
Whole vs. parts		1.00	*.09(<.001)**	*.15(<.001)**	*.07(<.001)**	*.01(.16)*
Same vs. different			1.00	*.22(<.001)**	*.23(<.001)**	*.008(.33)*
Concrete vs. abstract				1.00	*.20(<.001)**	*.07(<.001)**
Novel vs. familiar					1.00	*.02(.005)**
Toward vs. away						1.00