Semantic Memory and Creativity:

The Costs and Benefits of Semantic Memory Structure in Generating Original Ideas

Roger E. Beaty[1*], Yoed N. Kenett[2*], Richard W. Hass[3], & Daniel L. Schacter[4]

[1]Department of Psychology, Pennsylvania State University, University Park, PA

[2]William Davidson Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Israel

[3]Jefferson Center for Interprofessional Practice and Education, Thomas Jefferson University, Philadelphia, PA

[4]Department of Psychology, Harvard University, Cambridge, MA

[*]Authors contributed equally

Author Note

Correspondence should be addressed to Yoed Kenett, 521 Bloomfield building, Kiryat Hatechnion, Haifa 3200003, Israel, yoedk@technion.ac.il

Abstract

Despite its theoretical importance, little is known about how semantic memory structure facilitates and constrains creative idea production. We examine whether the semantic richness of a concept has both benefits and costs to creative idea production. Specifically, we tested whether cue set-size—an index of semantic richness reflecting the average number of elements associated with a given concept—impacts the quantity (fluency) and quality (originality) of responses generated during the alternate uses task (AUT). Across four studies, we show that low-association, sparse, AUT cues benefit originality at the cost of fluency compared to high-association, rich, AUT cues. Furthermore, we found an interaction with individual differences in fluid intelligence in the low-association AUT cues, suggesting that constraints of sparse semantic knowledge can be overcome with top-down intervention. The findings indicate that semantic richness differentially impacts the quality and quantity of generated ideas, and that cognitive control processes can facilitate idea production when conceptual knowledge is limited.

Keywords: creativity, divergent thinking, fluid intelligence, semantic memory, cue set-size

Semantic Memory and Creativity:

The Costs and Benefits of Semantic Memory Structure in Generating Original Ideas

Creativity theories have long emphasized the role of semantic memory for creative thought (Abraham, 2014; Kenett, 2018b; Kenett & Faust, 2019; Mednick, 1962; Sowden et al., 2014). Semantic memory stores facts, concepts, and general knowledge that can be retrieved and combined in new ways to facilitate creative thinking (Kumar, 2021). According to the associative theory of creativity (Mednick, 1962), creative thinking involves connecting concepts in memory that are weakly related, or "far away" from each other, in a novel and appropriate way—analogous to the process of spreading of activation through concepts in semantic memory (Collins & Loftus, 1975). Thus, as the semantic "distance" between two concepts increases, so does the likelihood of the combined concepts being deemed creative (Beaty & Johnson, 2021; Green, 2016; Kenett, 2018a, 2018b; Mednick, 1962).

Despite the theoretical relevance of semantic memory for creative thinking, little empirical work has examined its specific cognitive contributions to creativity (Benedek & Neubauer, 2013; Kenett & Faust, 2019; Volle, 2018), leaving questions about how and when this memory system impacts creative thought (Beaty et al., 2017; Kenett, 2018b). Although semantic memory is undoubtedly necessary for creative processes such as conceptual combination, problem solving, or imagination (Abraham, 2014; Abraham & Bubic, 2015; Bieth et al., 2021; Kenett & Thompson-Schill, 2020; Schilling, 2005), research has also documented constraints of semantic knowledge on creative performance such as functional fixedness—a biasing of attention to salient conceptual features that prevents cognitive flexibility (Chrysikou et al., 2016; Chrysikou & Weisberg, 2005). Extant findings thus point to potential benefits and costs of semantic memory to creative thinking. In the present research, we examine whether creative thinking is impacted by semantic memory structure—focusing on the richness of related concepts within semantic networks—with the aim of identifying potential costs and benefits of memory structure to creative idea production.

Creative Thinking and Semantic Memory

The associative theory of creativity (Mednick, 1962) proposes that creative individuals have a richer associative memory structure than less creative individuals. According to this theory (Mednick, 1962), creative individuals are characterized by ‘‘flat” (more and broader associations to a given concept) instead of “steep” associational hierarchies (few, common associations to a given concept). On this view, creative individuals may have more associative links between concepts in their semantic memory and can connect associative relations faster than less creative individuals, thereby facilitating more efficient search processes (Beaty et al., 2021; Gray et al., 2019; Kenett, 2018b; Kenett & Austerweil, 2016; Volle, 2018). Thus, when attempting to think creatively, a less creative individual is likely to become “stuck” on these dominant, common, associations whereas a more creative individual can overcome them and establish more distant associations, via spreading activation (Beaty et al., 2021; Gray et al., 2019; Kenett & Austerweil, 2016). Within this framework, semantically “close” concepts are considered less likely to be creative, whereas semantically “distant” concepts are often considered creative (Heinen & Johnson, 2018; Kenett, 2018a, 2019)[1].

While still debated (Benedek & Neubauer, 2013), the associative theory of creativity has received recent empirical support from computational investigations of individual differences in creative thinking (Benedek et al., 2017; He et al., 2021; Kenett et al., 2014; Kenett et al., 2016; Kenett et al., 2018; Ovando-Tellez et al., 2022). For example, Kenett and colleagues conducted a computational network analysis of free association data to compare the semantic memory network structure of low- and high-creative individuals, defined by performance on a battery of creative thinking tasks. The authors found that the semantic memory network of the high-creative group was characterized by higher connectivity and lower overall distances between concepts (Kenett et al., 2014), likely permitting more efficient spreading activation processes to unfold (Kenett et al., 2018). Importantly, these semantic memory network characteristics have been shown to facilitate broader search processes (Kenett & Austerweil, 2016) that allow reaching weaker, and more uncommon concepts (Stella & Kenett, 2019). These findings have since been replicated and extended in other group-based analyses (Beaty et al., 2021; Gray et al., 2019; Kenett et

[1] Creativity is often defined by a combination of novelty and appropriateness. Semantic distance captures the novelty aspect of creativity. Thus, two concepts that are semantically distant may be considered “creative” to the extent that the distant association is also appropriate.

al., 2016), as well as at the individual level (Benedek et al., 2017; He et al., 2021; Ovando-Tellez et al., 2022). A similar semantic memory network structure was also reported in a study on Openness to Experience—a personality trait linked to creative behavior and cognition (Christensen et al., 2018)—providing further support for the role of semantic memory in creativity (Kenett & Faust, 2019; Volle, 2018). Taken together, these studies highlight the role of *semantic richness* in creative thinking, suggesting that high connectivity between concepts supports creative idea production.

Effects of Semantic Richness on Retrieval Processes

Several studies have demonstrated a role for semantic richness in facilitating word recognition and retrieval. One measure of semantic richness relates to the number of semantic “neighbors” (or semantic density) linked to a given cue word (Al-Azary & Buchanan, 2017; Balota et al., 2004; Danguecan & Buchanan, 2016; Duñabeitia et al., 2008; Pexman et al., 2008; Recchia & Jones, 2012; Yap et al., 2012). Here, semantic richness is quantified either by the different number of unique associative responses generated to a cue word (Nelson et al., 2004) or the number of words that are generally “close” to the cue word in a semantic space (Pexman et al., 2008). Specifically, the cue set-size effect refers to cues that have a larger associative set (number of unique associative responses) impacts memory recall (Nelson & McEvoy, 1979; Nelson et al., 1999).

In a similar vein, Mak et al. examined the role of associative degree centrality in memory recall (Mak et al., 2021; Mak & Twitchell, 2020). Semantic degree centrality refers to the number of connections a word has in a semantic memory network, or space (Hills & Kenett, 2022). Thus, words with many, rich, connections are considered as high-degree words, whereas words with few, sparse, connections are considered as low-degree words. Mak et al. compute associative degree centrality based on De Deyne et al.’s English Small World of Words dataset, a large-scale of free-association responses collected from 88,722 participants to 12,292 cue words (De Deyne et al., 2019). Based on this dataset, a cue-word out-degree refers to the number of unique associative responses a cue word has; in-degree refers to the number of times that a cue word is generated as an associative response in the dataset. Thus, these measures correspond to the notions of forward and backward associative strength proposed by Nelson et al., reflected in the University of South Florida Free Association Norms (Nelson et al., 2004). Based on this associative degree centrality measure, Mak et al. show how it influences immediate serial recall as well as the learning of new words (Mak et al., 2021; Mak & Twitchell, 2020).

Richer semantic neighborhoods have been shown to facilitate a range of linguistic tasks, including word naming (Balota et al., 2004), visual word recognition, lexical decision tasks, word demasking (Duñabeitia et al., 2008), abstract word processing (Danguecan & Buchanan, 2016; Recchia & Jones, 2012) and metaphor comprehension (Al-Azary & Buchanan, 2017), and constraining retrieval processes (Marko & Riečanský, 2021). Overall, these consistent findings point to a benefit of a richer semantic neighborhood on word recognition and processing. The results can be explained by the spreading activation model proposed by Collins and Loftus (1975), which proposed that semantic memory is organized as a network, upon which cognitive activation spreads through it and decays rapidly over time and distance (Hills & Kenett, 2022; Kumar, 2021; Siew, 2019). In line with this theory, because of the bi-directional spread of the activation, cue words with richer semantic neighborhoods will receive more activation than words with sparser semantic neighborhoods. A similar theory is known as the connectivity model (Gentner, 1981; Klimesch, 1987; Kroll & Klimesch, 1992). According to this model, the greater the number of connected words a cue word has, the more efficient its processing and retrieval (see also Marko & Riečanský, 2021). Specifically, the connectivity model has four assumptions (Klimesch, 1987): 1) concepts in memory are represented in an interconnected structure (similar to Collins & Loftus, 1975); 2) activation spreads in a direct, but also indirect manner—as such, the richer the semantic neighborhood of a cue word is, the more indirect activation that can spread from the cue word in a forward and backward fashion; 3) a search process in semantic memory terminates when indirect activation spreads back to the originating cue word, and the search process will terminate if no indirect activation reaches the originating node after a given amount of time; 4) semantic processing time is a function of the amount of indirect activation. According to the connectivity model, concepts with richer semantic neighborhoods will be processed faster than concepts with sparser semantic neighborhoods (Kroll & Klimesch, 1992). Based on this model, and in line with the associative theory of creativity (Mednick, 1962), a richer semantic memory structure should benefit the production of original ideas.

Semantic Memory Constraints on Creative Thinking

Despite growing empirical support for the associative theory of creativity and the benefits of semantic memory structure to creative thinking, past work has also demonstrated constraining effects of semantic memory on creative performance. Such work includes fixating on stereotypical object information (i.e., functional fixedness; Chrysikou et al., 2016; Chrysikou & Weisberg, 2005; Glucksberg & Weisberg, 1966), biasing idea generation with salient examples (Beaty et al., 2017; Chrysikou et al., 2016; Marsh et al., 1996; Smith et al., 1993), and priming incorrect solutions on the remote associates test (Smith & Blankenship, 1991). Together, these studies indicate that activating highly related conceptual knowledge within semantic networks can interfere with spreading activation to remote concepts during creative thinking.

Additional studies have shown that rich knowledge can hinder creative thinking (Forthmann et al., 2016; Rietzschel et al., 2007; Wiley, 1998). Wiley (1998) examined how the knowledge structure of experts possibly constrains them and leads to fixation in creative problem solving; notably, Wiley (1998) examined the role of domain-specific knowledge in experts, whereas other work has focused on the role of domain-general semantic knowledge. Rich, structured knowledge in a domain allows an expert in that domain to be highly efficient in processing and retrieving information related to that domain. However, such a rich domain-based knowledge also leads to fixation and hinders creative problem solving. Across a series of studies, Wiley (1998) found that experts were more prone to fixations in misleading problems. Wiley attributes the effect of domain knowledge on experts to narrowing the scope of their search through it, thus preventing broad search processes (Wiley, 1998). Therefore, a rich body of knowledge can lead to benefits in one case, and constraints in another. Similarly, although a rich semantic neighborhood can facilitate word recognition and activate a broader set of associated responses, it may hinder moving farther away from the original concept to generate a novel and original idea (Kenett, 2018a; Mednick, 1962).

A few studies have empirically demonstrated a cost-benefit effect in generating responses in creativity tasks (Forthmann et al., 2016; Rietzschel et al., 2007). Forthmann et al. (2016) examined the effect of different types of instructions (be-fluent vs. be-creative) on participants performance in a divergent thinking task for low- and high- frequency words (directly related to the breadth of their semantic neighborhood; Cofer & Shevitz, 1952). Participants generated more responses for high-frequency words, but only in the be-fluent condition, whereas instruction type did not affect the number of responses generated for low-frequency words (Forthmann et al., 2016). The authors interpret their findings as participants needing to utilize cognitive control mechanisms to filter salient responses during the "be-creative" condition for high-frequency responses. Rietzschel, Nijstad, and Stroebe (2007) examined a common intuition in brainstorming research, i.e., that quantity and quality of ideas are related to each other. In a series of studies, the authors primed participants for deeper exploration of subcategories related to a problem, demonstrating its specific effect on the originality of ideas for that problem. Thus, the authors argue that fluency and originality of ideas have a complex relation that can be separately manipulated (Rietzschel et al., 2007). Overall, these studies highlight constraints of rich semantic neighborhoods, raising questions about whether and how such fixation effects can be overcome to generate original ideas.

Semantic Interference and Executive Control

Increasing evidence indicates that the creative thought process can be guided through executive control (Beaty et al., 2016; Benedek et al., 2014; Chrysikou, 2019; Nusbaum & Silvia, 2011). This work has highlighted the involvement of executive and strategic aspects of cognition required to control memory processes, including pre-potent response inhibition (Benedek et al., 2014), broad retrieval ability (Avitia & Kaufman, 2014; Silvia et al., 2013), and category switching (Nusbaum & Silvia, 2011). The joint contributions of spreading activation (bottom-up) and cognitive control mechanisms (top-down) are consistent with dual-process models of creative cognition (Barr et al., 2015; Kleinmintz et al., 2019; Sowden et al., 2014; Volle, 2018). In general, dual processes theories posit that creative thinking involves a dynamic interaction between idea generation and idea evaluation, corresponding to memory retrieval and cognitive control processes, respectively; notably, the extent to which these processes operate in serial or parallel remains unclear (Sowden et al., 2014).

To characterize the dynamics of creative thinking, researchers have examined a temporal trend in idea production known as the serial order effect (Acar et al., 2018; Beaty et al., 2014; Christensen et al., 1957; Hass & Beaty, 2018; Wang et al., 2017). The serial order effect is the tendency for responses to a divergent thinking task to become less frequent and more original over time. This temporal trend has been explained within the associative theory of creativity: at the beginning of an idea generation task (e.g., thinking of alternate uses for a common object), fluency is high because people typically begin by activating the rich semantic neighborhood surrounding the object cue, thus producing known uses for the object (Gilhooly et al., 2007; Matheson & Kenett, 2021). Likewise, fluency decreases and originality increases later in task because it takes time for spreading activation processes to unfold (i.e., to reach more distal concepts within a semantic network; Collins & Loftus, 1975; Matheson & Kenett, 2021; Mednick, 1962). Beaty and Silvia (2012) examined individual differences in the serial order effect and found that it interacted with fluid intelligence (*Gf*): as *Gf* increased, the serial order effect for originality diminished (see also Hass, 2017). Time was thus less relevant for originality at higher levels of intelligence, suggesting that cognitive control may mitigate early sources of semantic interference.

The Present Research

Past research has shed light on the contribution of semantic memory to creative cognition. At the individual level, for example, a more "flexible" (higher connectivity and shorter distances between concepts) semantic memory structure has been shown to facilitate creative idea generation (Beaty et al., 2014; Benedek et al., 2017; He et al., 2021; Kenett, 2019; Kenett et al., 2014; Ovando-Tellez et al., 2022). On the other hand, the activation of salient conceptual knowledge can constrain creative thought (Beaty et al., 2017; Smith & Blankenship, 1991). Together, these findings indicate that the organization of and access to concepts in semantic memory plays a key role in how they are retrieved during creative task performance. Furthermore, studies on other cognitive processes that are important in creative thinking, such as cognitive control and working memory capacity have shown both beneficial and constraining effects on creativity (Chrysikou, 2018; Van Stockum & DeCaro, 2020).

Does more or less knowledge about a concept influence how one thinks creatively with that concept? On the one hand, the associative theory of creativity posits that greater semantic richness facilitates spreading activation via more associative elements and connections (Mednick, 1962); in other words, the semantic infrastructure afforded by extensive concept knowledge makes spreading activation easier (Kenett & Faust, 2019; Volle, 2018). In the context of a divergent thinking task (or similar task involving idea generation), additional semantic knowledge may promote greater fluency as people “have more to say” about the concept. On the other hand, an executive processes interpretation views semantic richness as potentially detrimental to originality, because most closely-connected concepts are highly conceptually related and thus not novel (Beaty et al., 2014). However, such semantic interference can be overcome via suppressing its activation, as evidenced in studies of individual differences in executive control (Bunting et al., 2004). This work thus raises a fundamental question for creativity research; namely, *how and when does semantic memory structure facilitate and constrain creative thought?*

Our work is motivated by the fan effect. A fan effect is a memory phenomenon whereby increasing knowledge about a concept (or cue) leads to an increase in reaction time and accuracy on recognition memory tests, i.e., the more concepts “fanning” from a given cue, the more interference (Anderson, 1974; Anderson & Reder, 1999; Radvansky, 1999). The fan effect stems from the activation of competing concepts in memory. In the original work, Anderson (1974) showed that increasing the presentation of to-be-remembered items (i.e., propositional phrases) over the course of an encoding phase slowed reaction time (RT) and impaired accuracy in a subsequent recognition test. Anderson and colleagues have shown how the fan effect is found for the retrieval of schemas and real-world knowledge (Anderson & Reder, 1999), as well as the effect of prior knowledge on increasing interference during retrieval (Lewis & Anderson, 1976). Here, we conceptualized the fan effect as semantic richness, asking whether increased semantic richness interferes with divergent thinking as it does for other memory tasks.

We conducted a series of studies to test for the existence and impact of semantic richness in facilitating and constraining divergent thinking. To this end, we selected cue words (i.e., common objects) for the Alternate Uses Task (AUT; Acar & Runco, 2019; Kaufman et al., 2008; Torrance, 1972) that varied in semantic richness, defined as the average number of associations linked to the cue words based on free association norms (Nelson et al., 2004). We reasoned that, compared to low-association object cues, high-association cues may yield a greater number of ideas (i.e., higher fluency) because such cues are presumably embedded within more densely connected semantic neighborhoods. However, in light of past work on the interfering effects of close conceptual knowledge (Beaty et al., 2017), we expected that this fluency benefit may come at the cost of originality as these ideas are likely to be less semantically distant (Study 1). At the same time, the relative sparseness of low-association cue neighborhoods may hinder spreading activation to distal concepts in the absence of a robust semantic infrastructure. To probe order effects of idea generation as a function of associative set-size, we tracked the serial order of response production (Study 2). In addition, we examined whether people with higher cognitive control ability may be more immune to the potential set-size effect (Study 3). Finally, we assessed whether the effects of semantic richness could be detected using computational measures of idea originality based on semantic distance (Study 4).

Study 1

In Study 1, we aimed to test for the existence of a semantic richness effect in divergent thinking performance. To this end, we developed an experimental manipulation of cue type in the Alternative Uses Task (AUT), a commonly used assessment of divergent creative thinking (Acar & Runco, 2019; Kaufman et al., 2008; Runco & Acar, 2012). We identified AUT cues (i.e., common objects) that varied as a function of cue association set size—an index of semantic richness defined as the average number of free associations generated by participants in a widely used free association norms dataset (Nelson et al., 2004)—thus yielding high-association, rich, cues (i.e., rich semantic cues, *RSC*) and low-association, sparse, cues (i.e., sparse semantic cues, *SSC*). We hypothesized that, compared to SSC, participants would generate significantly more AUT responses to RSC due to increased semantic richness of associations related to these cues. Furthermore, we hypothesized that this fluency benefit would come at a cost of originality, such that participants would generate AUT responses rated as more original to SSC due to a decreased presence of salient and highly-conceptually related (i.e., unoriginal) concepts.

Method

Participants

Forty participants were recruited for the study via Amazon Mechanical Turk (AMT; Buhrmester et al., 2011). Participants were offered $3.00 compensation for completion of the entire study. No participants' work was rejected (i.e., all 40 participants were paid), however, a pre-analysis screening procedure identified 4 participants that failed to provide responses for all 10 cues and thus did not follow instructions. The final sample size for analysis was 36 participants (19 female) with an average age of 36.34 years (SD = 11.71 years).

Materials

Stimuli. We began by constructing a list of SSC and RSC for the AUT. SSC and RSC were selected from the University of South Florida Free Association Norms database, which includes norms for 5,018 cue words (Nelson et al., 2004). Importantly, for each of these cue words, the database lists the number and types of different associative responses that were generated to these cue words. The number of associative responses to a cue word was used as a proxy of the set size of the cue word. In addition, for each associative response to a cue word exists its forward and backward associative strengths (a responses out- and in-degree centalityMak et al., 2021). However, since we were interested in examining the effect of the general set-size of the cue on the AUT, we only consider the number of associative responses in the current study. Of the 5,018 cue words, we manually selected cue words of concrete objects that can be used in an AUT. Finally, a list of five SSC (clock, fork, lamp, lens, pen) and five RSC (soap, rope, stick, marble, balloon) were retained. These cue words were matched on key linguistic variables: *frequency* (RSC M = 21.4, SD = 10.97; SSC M = 16.4, SD = 3.29; $t(8) = 1.00$, $p = .35$) and *concreteness* (RSC M = 6.09, SD = .23; SSC M = 5.88, SD = .67; $t(8) = .66$, $p = .53$). Critically, the average set size of RSC cues (M = 22, SD = 1.22) was significantly greater than the average set size of the SSC cues (M = 6.6, SD = 1.51; $t(8) = 17.67$, $p < .001$).

AUT. For each of the ten cue words (SSC and RSC), participants had two minutes to generate as many alternate uses as possible, a time window typically used in AUT tasks (Acar & Runco, 2019). Two main measures were computed from participants AUT performance: originality (i.e., average subjective rating) and fluency (i.e., sum of responses). Originality was rated on a 5-point scale (ranging from *1 – Very obvious/ordinary use* to *5 – Very imaginative/re-contextualized use*) designed for cognitive studies of divergent thinking (Hass et al., 2018). Responses were rated by three AMT participants not involved in the experiment. Raters rated an alphabetized set of unique responses per prompt, and were not aware of any order, or which participants produced those responses.

Procedure

The AUT was administered online via Qualtrics (www.qualtrics.com). Participants were initially given the following instructions: "We want you to come up with as many original and creative uses to objects as you can. The goal is to come up with creative ideas, which are ideas that strike people as clever, unusual, interesting, uncommon, humorous, innovative or different. You will do this for ten different objects, with two minutes for each object". RSC trials (n = 5) were completed in one block and SSC trials (n = 5) were completed in a separate block; the order of the blocks was counterbalanced across participants. The Qualtrics interface consisted of an instructions page and a response-collection page. For each AUT trial, participants had two minutes to type as many alternate uses as possible into textboxes presented on the response-collection page. The Study did not include initial practice, and participants were given a short break between blocks. Following completion of the AUT blocks, participants completed a short demographic survey.

Results

Inter-rater reliability was assessed with intraclass correlations ICC(2,3), and was generally high across the 10 cues (M = .61, SD = .11).

To test whether AUT performance varied as a function of set size, we contrasted fluency and originality of AUT responses across the SSC and RSC conditions by computing paired t-test analyses on the means.

Regarding fluency, participants generated significantly more responses in the RSC condition (M = 5.93, SD = 2.3) compared to the SSC condition (M = 5.26, SD = 1.85; $t(35) = 4.13$, $p < .001$, d = .69, 95% CI = [.34, 1]). Regarding originality, participants generated significantly less original responses in the RSC condition (M = 2.50, SD = .30) compared to the SSC condition (M = 2.72, SD = .34; $t(35) = -5.95$, $p < .001$, d = 1.2, 95% CI = [-.27, -.15]). Thus, the AUT responses to RSC yielded higher fluency but lower originality when compared to the AUT responses to SSC.

Discussion

The associative theory of creativity implicates spreading activation of concepts within semantic memory, but little is yet known about the benefits—and potential costs—of semantic memory in creative thinking. Study 1 identified one such benefit and cost of semantic knowledge to performance on the AUT. Participants generated more responses during the AUT when using RSC compared to SSC, suggesting that greater semantic content benefits ideational fluency. This benefit, however, came at the cost of originality: participants generated ideas that were rated as less creative in the RSC condition. This finding is consistent with the notion that salient conceptual information (e.g., RSC associations) can constrain creative thought by acting as a source of interference that must be inhibited to establish more remote conceptual combinations (Beaty et al., 2017; Chrysikou, 2019).

Study 2

The goal of Study 1 was to obtain preliminary evidence for the existence of a semantic set size effect on the AUT. In Study 2, we sought to replicate this effect in a larger sample of participants using a new online platform with greater experimental control than Study 1 (Hass & Beaty, 2018). This platform allowed us to further probe potential order effects of the set-size effect. Specifically, we examined whether the set-size effect interacted with the serial order effect—the tendency for AUT response originality to increase over time (Bai et al., 2021; Beaty & Silvia, 2012).

Method

Participants

Fifty-five participants were recruited for the study via AMT. We built on the sample size of Study 1 as a benchmark (*n* = 40) for replication and increased the sample size based on the availability of funds. Participants were offered $4.00 US compensation for completion of the study. No participants' work was rejected (i.e., all 55 participants were paid), however, a pre-analysis screening procedure identified 14 participants that failed to provide responses for all 10 cue words and 1 participant that provided random responses and thus did not follow directions. The final sample size for analysis was 40 participants (30 female) with an average age of 38.1 years (SD = 12.07 years). This study was approved by Jefferson University's institutional review board.

Materials

Stimuli. The same stimuli used in Study 1 were used here (10 AUT cues; 5 SSC, 5 RSC).

Divergent Thinking Task. For each of the ten cue words (SSC and RSC), participants had three minutes to generate as many alternate uses as possible. Time window was extended from two minutes (as used in Study 1) to three minutes per object so to have more time for the serial order effect to unfold (Beaty & Silvia, 2012). Like Study 1, fluency and originality were computed for each participant. Furthermore, we also logged *inter-response time* (the time between the first key strokes of successive responses) and the *order* of entry of each response for the serial order analyses.

Using the tools created by the psiTturk project (Gureckis et al., 2016), a custom web application was employed for administering the experimental tasks (for details on the platform, see Hass & Beaty, 2018). Similar to Study 1, the interface consisted of an instructions page and a response-collection interface. The instructions page appeared before both blocks of trials (SSC and RSC); after reading instructions, participants moved on to the tasks using a navigation button. The task interface consisted of a text-display, which contained the object prompt for that trial and a text-entry field. JavaScript code saved the first key press per response, the time at which the participant entered the response (by pressing ENTER or RETURN), and the text of the response itself.

Procedure

Upon accepting the HIT on Amazon Mechanical Turk (Buhrmester et al., 2011), the psiTurk tools generated the experimental environment. Following consent and instructions, participants completed a practice trial to acclimate to the typed-entry interface, which involved typing the names of colors that they knew for 30 seconds. Upon completion of practice, the first set of experimental trials began. Participants were informed that there would be five trials, each with a different object, and each lasting 3 minutes. SSC and RSC conditions were separated by another instruction page which simply reiterated the previous instructions and informed participants that they could take a short break. The order of blocks and cues within blocks was randomized. Following the AUT, participants completed a short demographic questionnaire, including their level of engagement with the experiment.

Results

Participants' responses were rated for originality on the same 5-point scale as Study 1 (Hass et al., 2018) by two research assistants and one AMT worker not involved in the experiment. Inter-rater reliability ICC(2,3) ranged from fair to good across the 10 cues (M = .47, SD = .15).

Participant-level Fluency and Originality

We first attempted to replicate the results of Study 1, assessing whether fluency and originality varied as a function of set-size. Regarding fluency, participants generated significantly more responses in the RSC condition (M = 9.17, SD = 3.42) compared to the SSC condition (M = 8.10, SD = 3.10; $t(39) = 3.84$, $p < .001$, d = 0.61, 95% CI = [0.51, 1.64]). Regarding originality, participants generated significantly less original responses in the RSC condition (M = 2.67, SD = .26) compared to the SSC condition (M = 3.03, SD = .26 $t(39) = 6.47$, $p < .001$, d = 1.02, 95% CI = [0.18, 0.36]). These results replicate the findings of Study 1 and suggest that a higher number of associative links afforded by RSC benefits fluency at the cost of originality, potentially due to increased interference from salient but semantically similar concepts. Note that mean fluency values were higher in Study 2 compared to Study 1 due to longer trials (2 min vs. 3 min) for response-level analysis.

Response-level Effects: IRT and Serial Order

To further investigate potential effects of the set-size manipulation, two response-level analyses were performed. First, inter-response times (IRTs) were compared across the two conditions with a mixed-effects regression model. To conform to model assumptions (namely normally distributed residuals), IRTs were log-transformed and regressed on 1) a fixed-effect of condition (SSC vs. RSC), 2) a random effect of participant, and 3) a random effect of prompt. Though mean IRTs were shorter in the RSC condition (M = 14.50s, SD = 13.99s) compared with the SSC condition (M = 16.14s, SD = 16.63s), the fixed effect in the log-IRT model was not significant, b = .0004, $p = .55$.

Next, we examined the relationship between response order and originality rating per cue with a mixed-effects model. The model included a linear serial order term that was scaled so that the first response was equal to zero. This yields an interpretation of the intercept as the mean originality of the first response for the RSC condition, and an interpretation of an effect of set-size as the difference in originality of the first responses between conditions. An interaction between condition (SSC vs. RSC) and serial order was also modeled, along with random effects of participant and prompt. The full model results are presented in Table 1. There was a significant serial order effect, b = 0.016, $p < .001$; but the overall difference between SSC and RSC originality was not preserved in this model, b = 0.077, $p = .52$. Additionally, there was no significant difference between the conditions in terms of the linear slope, b = 0.011, $p = .08$.

Table 1: Linear mixed effect model of originality in study 2

Fixed Effects	*B*	SE	*p*
Intercept	2.07	0.088	< .001
Serial Order	0.02	0.004	< .001
Set-size	0.08	0.116	.52
Serial Order*Set-size	0.01	0.006	.08
Random Effects	Name	Variance	SD
Participant	Intercept	0.05	0.21
Cue	Intercept	0.03	0.17
Residual		0.46	0.68

Full model: Originality ~ Set-size + (order – 1) + (order -1)*Set-size + (1|participant) + (1|cue)

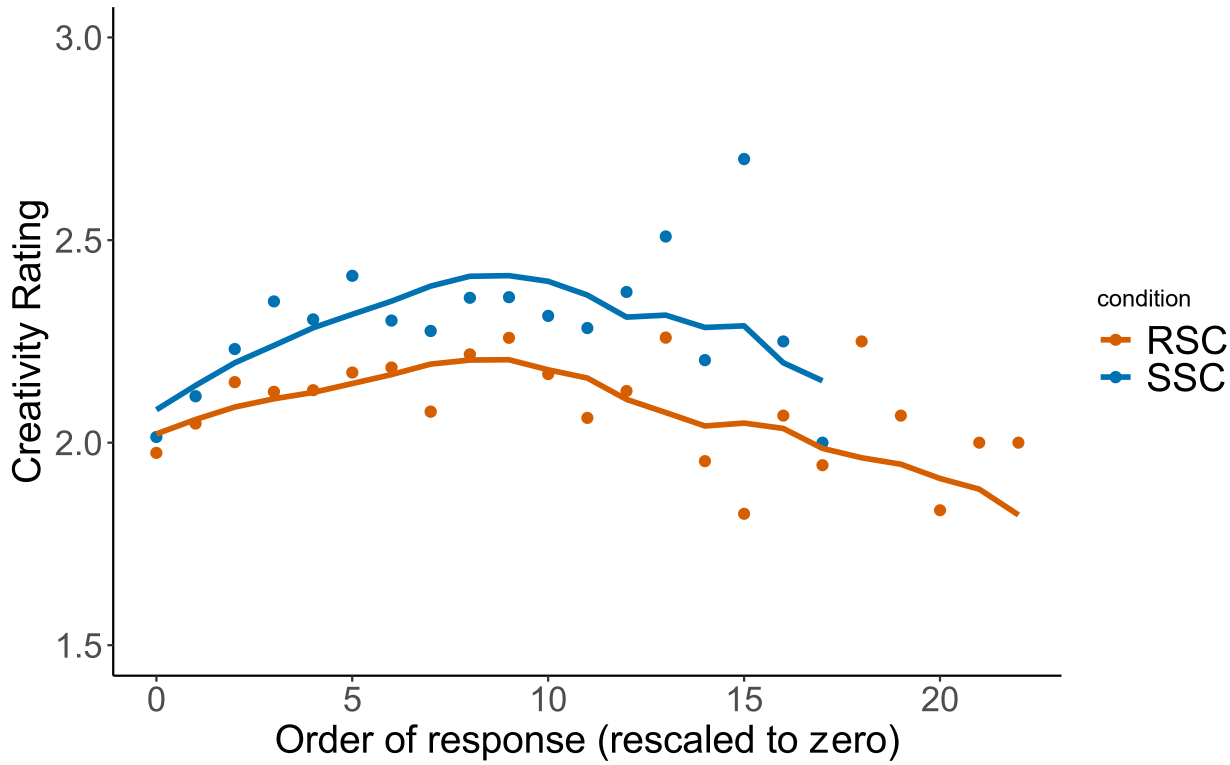

Fig. 1: Average originality rating as a function of response order. The first response was rescaled to zero for the purpose of growth modeling. Lines represent fitted values for each condition based on the multilevel linear growth model.

Discussion

The aim of Study 2 was to replicate the results of Study 1 in a larger sample and to examine the interaction between the serial order effect (Beaty & Silvia, 2012) and the set-size effect. Study 2 replicated the results of Study 1: The RSC condition led to increased fluency at a cost of lower originality. While we found a main effect of serial order on response originality, in line with previous studies, we did not find a significant interaction between serial order and set-size effect. A potential cause for this null finding relates to the effect size: the mean originality of SSC responses is clearly higher for most responses after the 3rd response, but the effect is on the order of about a quarter of a point on the 5-point rating scale. However, when generating the first few SSC or RSC responses, originality seems to vary less. This pattern may be a function of the fact that there are far more data points per participant in these early phases of the experiment, or it may be that set-size has more of an effect when participants have exhausted a few, potentially welllearned responses. When responses are aggregated at the level of participant, these serial order effects are seemingly eliminated, leading to an overall significant uptick in originality to SSC responses.

Study 3

Studies 1 & 2 established a semantic set-size effect on the AUT. In Study 3, we sought to further replicate and extend this finding. Specifically, we employed the same experimental paradigm—varying cue set size across AUT items—and further examined potential interactions with *Gf*, an individual difference variable with established links to divergent thinking (Bai et al., 2021; Beaty & Silvia, 2012). Although *Gf* has been shown to predict the creative quality of AUT responses, the cognitive contribution of *Gf* to creative performance remains largely unknown. One possibility is that *Gf* supports inhibitory control processes, consistent with its strong association with executive control (Kane et al., 2005). Thus, *Gf* may be more relevant for RSC idea generation via the inhibition of salient conceptual knowledge (Beaty & Silvia, 2012). On the other hand, *Gf* may support SSC idea generation by facilitating spreading activation within a relatively sparse semantic neighborhood. In addition to examining the role of *Gf*, we further probed order effects of the set-size effect (i.e., serial order) as a function of cue set size.

Method

Participants

One hundred thirteen participants (50 females) were recruited from AMT. Sample size was based on similar previous studies of divergent thinking and *Gf* (Beaty & Silvia, 2012; Benedek et al., 2017). The average age of participants was 37.71 years (SD = 10.49). Participants received $5.50 for completion of the experiment. Thirty-three participants were excluded from the analysis due to failure to successfully complete all tasks or providing nonsensical answers to the open-ended questions. The final sample size for the current analysis was 83 participants (41 female, 1 prefer-not-to-answer) with an average age of 36.46 years (SD = 9.80 years). This study was approved by the Thomas Jefferson University institutional review board.

Materials

Stimuli. The stimuli used in Study 3 were identical to those used in Study 1.

AUT. The AUT used in Study 3 was identical to that used in Study 1. Specifically, participants had two minutes for each cue to generate as many alternate uses they could think of for that cue; AUT trial duration was two minutes to reduce the burden on participants completing 10 AUT trials and 3 intelligence tests. The software used to run the task was slightly modified to increase RT precision, i.e., a cue word was not displayed until the participant pressed the spacebar.

Fluid Intelligence. Based on Kenett et al. (2016), *Gf* was assessed via three separate tasks: 1) The series task from the Culture Fair Intelligence Test (CFIT), which involves choosing an image that correctly completes a series of images (13 items, 3 min); 2) A letter-sets task, which presents a series of four-letter combinations and requires people to determine which set does not follow a rule governing the other four (16 items, 4 min); and 3) A number-series task, which presents a sequence of numbers and requires participants to discover a rule governing their change (15 items, 5 min). To compute a general composite *Gf* score, we used principal component analysis, by summing the multiplication of each independent *Gf* score by its weight of the first unrotated principal component (Kenett et al., 2016).

Procedure

The procedure was similar to Study 2 in that data were collected using a custom psiTurk interface. Participants completed the AUT first, and then psiTurk linked to the *Gf* tasks, which were hosted via Qualtrics. Upon providing electronic consent, participants were presented with an overall description of their tasks: that they would be prompted to generate ideas about specific cues and then complete some cognitive tasks. Participants then completed a practice idea-generation trial to become acclimated to the typed entry interface (naming colors). Upon completion of practice, the first set of experimental trials began. The order of cues within blocks and block presentation were randomized, and participants had a short break between blocks. In addition to the *Gf* tasks, participants completed a short demographic survey. Three raters were trained to score responses for originality using the same 5-point originality scale used in Studies 1 and 2 (Hass et al., 2018).

Results

Inter-rater reliability was assessed with intraclass correlations ICC(2,3), and was generally high across the 10 cues (mean = .68, SD = .12).

Analyzing the fluency and originality of participants' responses, the results replicated findings from Studies 1 and 2: participants generated a significantly higher number of responses to RSC (M = 7.56, SD = 3.82) than to SSC (M = 6.33, SD = 3.04), $t(82) = -4.65$, $p < .001$, d = .51, 95% CI = [-1.75, -.70]. Furthermore, RSC responses were rated significantly less original (M = 3.04, SD = .33) compared with SSC responses (M = 3.12, SD = .44), $t(82) = 2.14$, $p = .035$, d = .23, 95% CI = [.01, .14].

Next, the relationship between response order and originality rating was examined via a mixed-effects model. Before computing the compiled *Gf* score, we examine the zero-lag correlations between the three *Gf* tasks that we use (**Table 2**). In our full model, *Gf*, set-size, and serial order were assigned as independent measures, and the originality ratings as the dependent measure. Interactions between set-size and *Gf*, interaction between set-size and serial order, and interaction between *Gf* and a linear serial order term was also modeled, along with random effects of participant and cue (**Table 3**).

Table 2: correlation analysis between all *Gf* measures

	CFIT	Letters	Number-set
CFIT	-	$.51^{**}$	$.55^{**}$
Letters		-	$.54^{**}$
Number-set			-

Note - ** - $p < .01$

We first compared this model to a model that only included the random effects and found that this model improved the fit to originality ratings, χ^2 (6, N = 83) = 105.52, $p < .001$. Specifically, we found a significant positive relation between each of the three main variables (*Gf*, set-size, and order) on participants' originality scores. Thus, we replicate and extend the results found in Study 1, and replicate previous findings on the effect of *Gf* on the AUT (Beaty & Silvia, 2012). Regarding the interaction terms, we found significant negative relations between both interaction terms (*Gf**set-size and order*set-size) on participants' originality scores (**Fig. 2**). However, due to high collinearity between the serial order variable and the interaction of *Gf* and serial order variable ($r = -.71$), the interaction effect of serial order and *Gf* was not significantly related to originality scores in this model.

Table 3: Linear mixed effect model of originality in study 3

Fixed Effects	*B*	SE	*p*
Intercept	2.28	0.18	< .001
Gf	0.05	0.01	< .001
Set-size	0.19	0.10	.05
Order	0.05	0.01	< .001
*Gf**Set-size	-0.02	0.00	< .001
Order*Set-size	-0.02	0.01	< .001

Random Effects	Name	Variance	SD
Participant	Intercept	0.09	0.30
Cue	Intercept	0.01	0.09
Residual		0.65	0.80

Full model: Originality ~ *Gf* + Set-size + order + *Gf**Set-size + order*Set-size + *Gf**order + (1|participant) + (1|cue)

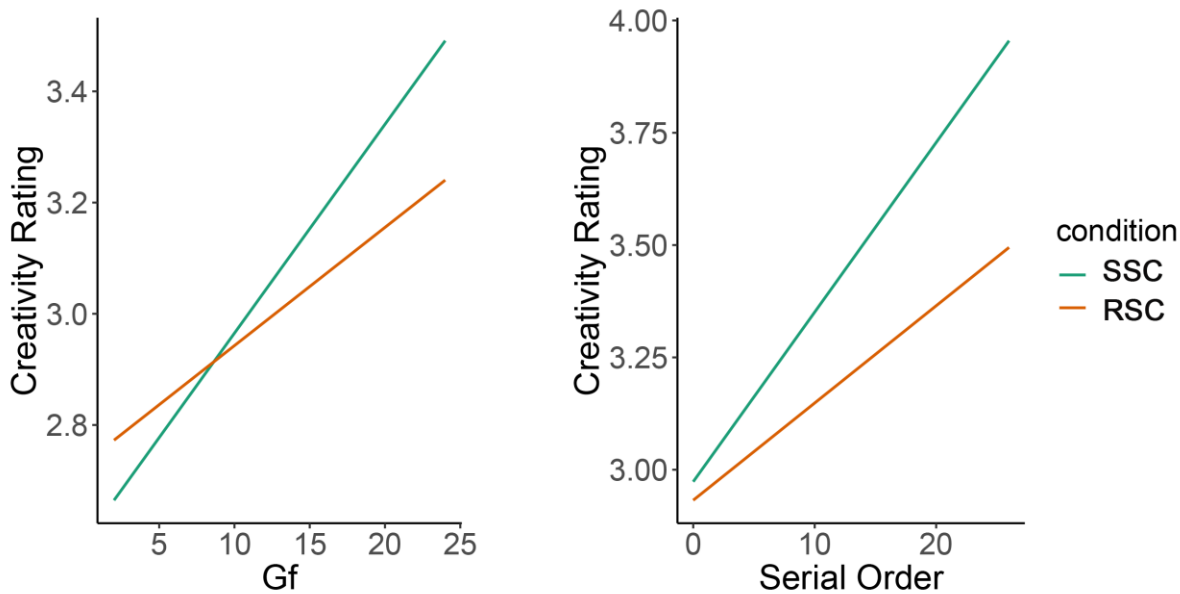

Fig. 2: Interaction effects between *Gf* and set-size effect (left) and serial order and set-size effect (right) on participant's originality ratings of their AUT responses.

Discussion

Study 3 replicated the findings of Studies 1 & 2, and extended them by examining individual differences in *Gf* (Beaty & Silvia, 2012). As in Studies 1 & 2, we found that, compared to SSC, RSC yielded increased fluency but decreased originality in the AUT. Study 3 further examined order effects associated with this set-size effect. Specifically, we replicated the serial order effect on the AUT—the tendency of idea originality to increase over time (Hass & Beaty, 2018)—and showed how this serial order effect interacted with both set-size (unlike in Study 2) and *Gf*. The ability to detect a significant interaction effect between serial effect and cue set-size in Study 3, unlike Study 2, might indicate that Study 2 was underpowered with regard to sample size (36 participants in Study 2 vs. 83 participants in Study 3). Although the 3-way interaction between serial order, set-size, and *Gf* was not significant due to exceedingly high collinearity between these independent variables, we found that interaction of *Gf**set-size and Order*set-size explained significant variance in originality ratings.

Study 4

Studies 1-3 established a semantic set-size effect on the AUT, highlighting a trade-off between fluency and originality of responses in relation to SSC versus RSC. Notably, however, all three studies used the same set of cue words as AUT cues, leaving questions regarding generalizability beyond this one list. Moreover, AUT responses were all scored for originality by human raters using the subjective scoring method (Hass et al., 2018). One possibility is that human raters simply found the responses to SSC more original due to particular affordances of the cues. In addition, it is possible that the originality effects reported in Studies 1-3 are confounded with fluency: if fluency differs across conditions, but the number of highly original ideas stays the same, then it will appear that originality has decreased in any condition where more ideas are generated, so the originality effect could be an artifact of the decision to average across ratings. In Study 4, we therefore aimed to replicate and extend our first three studies using 1) a different set of cue words; 2) an objective measure of originality based on computational assessments of semantic distance—a method increasingly used in creativity research to objectively quantify the novelty of responses

(Beaty & Johnson, 2021); and 3) an additional originality calculation that holds fluency constant across conditions. Lastly, we conducted this study in a controlled laboratory context, extending our first three studies which were conducted online.

Method

Participants

One hundred participants (55 females) were recruited from Pennsylvania State University. We roughly doubled the sample size of Study 1 and 2 to increase power and account for fewer available cues (see Materials). The average age of participants was 19.55 years (SD = 2.72 years). Participants received credit toward a course research option for their completion of the study. Seventeen participants were excluded from the analysis due to a failure to successfully complete all tasks or providing nonsensical answers to the open-ended questions. The final sample size for the current analysis was 83 participants (45 female) with an average age of 19.64 years (SD = 2.89 years). This study was approved by Pennsylvania State University's Institutional Review Board.

Materials

Stimuli. SSC and RSC stimuli were constructed similar to Study 1. Specifically, we manually searched through the University of South Florida Free Association Norms database (Nelson et al., 2004) for additional object cues that could be used on the AUT, carefully matching as closely as possible the linguistic properties of the stimuli used in Studies 1-3. This search resulted in a list of four SSC (brush, hammer, mirror, umbrella) and four RSC (barrel, basket, football, pants); note that Studies 1-3 used five cues per condition, hence the increase in sample size to boost statistical power (see Participants). These cue words were matched on key linguistic variables: *frequency* (RSC M = 21.5, SD = 11.45; SSC M = 22, SD = 17.07; $t(6) = 0.05$, $p = .96$) and *concreteness* (RSC M = 5.98, SD = .18; SSC M = 6.09, SD = .61; $t(6) = .32$, $p = .76$). Critically, the average set size of the RSC (M = 22, SD = 2.5) was significantly greater than the average set size of the SSC (M = 7.3, SD = .95; $t(6) = 11.21$, $p < .001$).

AUT. The AUT used in Study 4 was identical to Study 1 in terms of instructions and duration (2 minutes per cue), with the exception of the new cue words. Responses were objectively scored for originality using *SemDis*—an online platform for computing semantic distance (semdis.wlu.psu.edu; Beaty & Johnson, 2021). Semantic distance captures the novelty/originality dimension of creativity using computational models to quantify the relatedness between words in large corpora of natural language (Heinen & Johnson, 2018; Kenett, 2019). The semantic distance approach is increasingly employed in creativity research to automate originality scoring, with documented evidence for its validity, including high correlations with human creativity ratings (Beaty & Johnson, 2021; Dumas et al., 2021) and moderate correlations with other creativity measures (Beaty & Johnson, 2021; Prabhakaran et al., 2014). Consistent with applications in computational linguistics, semantic models are computed on word vectors within a high-dimensional space (Günther et al., 2019; Mandera et al., 2017).

Here, we leveraged the five semantic spaces available on the *SemDis* platform: a) two "count" models that count the co-occurrences of words within documents (Latent Semantic Analysis, LSA; global vectors, GloVe) and b) three continuous bag of words (CBOW) "predict" models that use neural network architectures with a sliding window to predict words given their surrounding context words (Mandera et al., 2017). The five models were built on large text corpora, including the 2009 Wikipedia dump (~800 million tokens), the ukwac web crawling corpus (~2 billion tokens), and the subtitle corpus (~385 million words; see Beaty & Johnson, 2021). A strength of using multiple semantic models is that it can mitigate the biases introduced by any single model or text corpus (Kenett, 2019). We used the following settings in *SemDis* to compute semantic distance: remove filler and clean, all semantic spaces, multiplicative compositional model (Beaty & Johnson, 2021). SemDis computes the semantic similarity between a given item (AUT object) and response (AUT response), then subtracts this similarity value from 1 to obtain a measure of semantic distance (Prabhakaran et al., 2014). For each of the eight AUT items, we averaged the semantic distance values for all five models for analysis. In addition, to control for the potential confounding effect of fluency across conditions, we used the "max-2" scoring method, where only the two most original responses for each participant/condition are averaged and included in the analysis.

Procedure

Participants completed the study in small groups (up to 5) at a testing laboratory at Pennsylvania State University. The lab is equipped with private testing cubicles running Windows desktop computers (Lenovo). An experimenter greeted participants and provided a brief explanation of the study. Following informed consent, participants completed the AUT and demographic information using the PsychoPy 3 experimental software (Peirce et al., 2019). Similar to Studies 1-3, following an instructions screen, participants completed a practice idea-generation trial. Upon completion of practice, the first set of experimental trials began. The order of cues within blocks and block presentation were randomized, and participants had a short break between blocks. The study took approximately 30 minutes to complete.

Results

Our first analysis aimed to replicate the set-size effect on AUT for fluency found in Studies 1-3 using the new list of AUT items. Paired-sample *t*-tests confirmed this replication: participants generated a significantly higher number of responses to RSC (M = 6.97, SD = 2.71) than to SSC (M = 6.44, SD = 2.10), $t(82) = -2.35$, $p = .021$, d = .26, 95% CI = [-.97, -.08].

Next, we assessed the effect of originality as objectively quantified by the composite average of the five semantic distance values. Consistent with the human-rated originality ratings reported in Studies 1-3, RSC responses yielded significantly lower semantic distance values (M = 0.92, SD = .05) compared with SSC responses (M = 0.94, SD = .04), $t(82) = 3.36$, $p < .001$, d = .37, 95% CI = [.01, .03]. Finally, we employed max-2 scoring—selecting and averaging the two most original/semantically-distant ideas per participant and condition—to assess whether the originality effect is robust to fluency. Critically, we found that max-2 semantic distance was significantly greater in the SSC condition (M = 1.02, SD = .03) compared to the RSC condition (M = 1, SD = .04): t(78) = -2.37, p = .02, d = .28, 95% CI = [-.02 -.01].

Discussion

Study 4 replicated and generalized the findings of Studies 1-3. Specifically, we used a new list of AUT items and an objective assessment of originality using computational models of semantic distance;

we also extended our first three online studies in a controlled lab context. Replicating prior results, compared to SSC AUT cues, participants generated significantly more AUT responses for RSC (higher fluency) but these responses were objectively quantified as less semantically-distant (lower originality), further suggesting a cost and benefit of semantic memory structure to divergent thinking.

General Discussion

Creative thinking has long been thought to benefit from associative processes unfolding in semantic memory through spreading activation (Kenett & Faust, 2019; Mednick, 1962). Here, we identify both benefits and costs of semantic memory for creative thinking, finding that increased semantic knowledge can benefit idea quantity at the cost of idea quality. We also show that the underlying structure of semantic knowledge interacts with cognitive control processes, extending recent work on the role of cognitive control in creative thought by identifying a key control mechanism that can strategically drive spreading activation in sparse semantic memory networks. Taken together, our findings provide insight into the role of, and interaction between, semantic memory and cognitive control during creative thought. Across four studies we demonstrate how cue set-size—an index of semantic richness reflecting the average number of elements associated with a given concept—differentially impacts the quality and quantity of divergent thinking responses.

Our findings indicate that increasing associations can cause a semantic set-size effect on AUT characterized by both a cost (decreased idea quality) and a benefit (increased idea quantity). In Study 1, we found that although participants generated significantly more AUT responses to rich-semantic cues compared to sparse-semantic cues (i.e., increased fluency), these responses were rated as significantly less creative (i.e., decreased originality). In Study 2, we replicated these results and also examined set-size effects in relation to the established serial order effect. While we found an effect of serial order, similar to previous studies (Bai et al., 2021; Beaty & Silvia, 2012), we did not find an interaction between serial order and set-size effect. This lack of interaction may be due to underpowered sample size, or additional cognitive processes, such as cognitive control (Chrysikou, 2019; Volle, 2018), but such control processes were not assessed in this study. In Study 3, we replicated the findings of Studies 1 and 2, and extended them by assessing individual differences in cognitive control (i.e., *Gf*) over a larger sample. We found that the set-size effect for originality varied as a function of *Gf*: as *Gf* increased, so did originality ratings in the SSC condition compared to the RSC condition. In addition, Study 3 replicated the serial order effect on AUT and found a significant interaction between serial order and set-size effects on the AUT. This result strengthens the trending interaction effect found in Study 2 (p = .08). In Study 4, we replicated and generalized the findings from Studies 1-3. We find the same fluency-originality tradeoff for RSC, using different lists of SSC and RSC words. In addition, we used an objective measure of originality, based on the quantitative semantic distance between the cue word and participants responses' (Beaty & Johnson, 2021). Taken together, the results extend recent work on the dynamics of memory retrieval and cognitive control during creative idea production (Benedek & Fink, 2019; Chrysikou, 2019; Volle, 2018).

Cost and Benefits of Semantic Memory

The current findings have implications for the associative theory of creativity (Mednick, 1962). According to the associative theory, creative thinking involves creating connections between concepts stored in semantic memory, and individual differences in creative thinking ability can be explained by variation in the organization of concepts. The theoretical work of Mednick (1962) has since received empirical support from several studies using computational network science methods to quantify semantic networks in low and high creative individuals (Benedek et al., 2017; He et al., 2021; Kenett et al., 2016; Kenett & Faust, 2019; Ovando-Tellez et al., 2022; Stella & Kenett, 2019), finding that high creative individuals have a more "flexible" semantic network structure—higher connectivity and shorter distances between concepts in these networks—conducive to remote conceptual combination. Our findings are consistent with the associative theory of creativity: on the one hand, a rich sematic neighborhood can benefit creative thinking by providing more associative links/ideas (high ideational fluency), facilitating the spread of activation (Kenett & Faust, 2019); on the other hand, a rich semantic neighborhood comes at the cost of increased interference (lower ideational originality), potentially leading to higher semantic fixation (Beaty et al., 2017).

Notably, Mednick (1962) emphasized the importance of semantic network structure for creative thinking, but he did not account for cognitive factors that can operate on this structure, such as cognitive control (Chrysikou, 2019; Volle, 2018). Consistent with past work (Beaty & Silvia, 2012; Benedek et al., 2014), Study 3 found that *Gf* predicted the originality quality of AUT responses. Critically, we found that *Gf* interacted with the set-size effect: higher-*Gf* benefited originality with SSC. From a semantic network perspective, SSC may be embedded in a less densely connected semantic neighborhood, potentially blunting spreading activation to remote concepts due to less semantic scaffolding (Kenett, 2018b; Klimesch, 1987; Mednick, 1962). Thus, one possibility is that *Gf* compensates for such sparse semantic connectivity by driving search processes in a top-down fashion (Volle, 2018). In other words, when less is known about an object, cognitive control may facilitate strategic and deliberate conceptual combination.

On the other hand, one might predict *Gf* to benefit RSC originality. Because RSC are likely embedded within a relatively denser neighborhood of semantic associations—as reflected by higher ideational fluency in the RSC condition across all studies—these associations may have induced interference due to high salience and semantic relatedness. Prior research suggests that salient concepts can disrupt idea generation by priming what is already known and thus not original (Beaty et al., 2017; Chrysikou et al., 2016). Thus, cognitive control could benefit RSC via inhibitory mechanisms, i.e., suppressing dominant responses and redirecting search processes (Beaty & Silvia, 2012). Notably, Study 3 assessed *Gf*—a proxy measure of cognitive control that strongly correlates with executive processes, such as inhibitory control (Kane et al., 2005). Future work might resolve this question by examining the contribution of specific executive functions to idea generation under similar semantic constraints.

More generally, our findings on a cost-benefit effect of set-size on AUT responses extends previous findings on linguistic "neighborhood" (phonological, orthographic, semantic) effects (Klimesch, 1987; Kroll & Klimesch, 1992; Luce & Pisoni, 1998; Mirman & Magnuson, 2008; Pexman et al., 2008). For example, Mirman and Magnuson (2008) found opposing effects of semantic neighbors on word recognition: near neighbors inhibited word recognition, while distant neighbors facilitated word recognition. The authors argue that this opposing effect is due an attractor dynamic effect: near neighbors act as competing attractors that inhibit word recognition, while distant neighbors create a gradient that facilitates settling on the correct (recognized word) attractor (Mirman & Magnuson, 2008). According to the connectivity hypothesis (Klimesch, 1987), an active search through semantic memory is terminated when indirect activity gets back to the original cue. On this view, RSC—embedded in a richer semantic neighborhood with denser connectivity—is more conducive for indirect activity to reach the original cue, simply by the random nature of the spreading activation process (Collins & Loftus, 1975).

Related to this speculation, empirical work analyzing phonological and semantic networks has demonstrated how richer neighborhoods may “trap” activation in them (Kenett et al., 2018; Siew, 2013), activation that theoretically decays rapidly over time. Such an explanation may also be related to the findings of Wiley (1998) who argues that experts have a narrower search process throughout the domain of their expertise. Thus, while RSC activate a larger number of associated concepts (higher fluency), these concepts may be more prototypical and salient (lower originality). As proposed by the associative theory of creativity, creativity is related to the ability to move farther away from a concept in memory (Kenett, 2018a; Kenett & Faust, 2019; Mednick, 1962). Here, we show that spreading activation during the creative thought process is constrained by the underlying structure of semantic knowledge.

Summary, Limitations, and Future Directions

The present research has potential implications for understanding the role of semantic knowledge in creative cognition (Kenett & Faust, 2019). Across four studies, we found a dissociation between the quantity and quality of ideas as a function of set-size: more ideas are generated when more was “known” about an object—as indexed via semantic associations—but these ideas were deemed to be of less creative quality. An interesting direction for future research would be to explore the extent to which this effect extends beyond “domain-general” creative performance to specific creative domains (cf. Wiley, 1998). Another outstanding question concerns whether the organization of semantic knowledge can be optimized for creativity through learning. We suspect that high creative ability is characterized by extensive domain-relevant knowledge, and superior access to that knowledge, via its hierarchical organization and top-down retrieval. A final future question concerns identifying the optimal density of semantic memory structure that is conducive to creative thinking—facilitating both quantity and quality of creative responses (cf. Faust & Kenett, 2014)—and determining how this optimal density varies across individuals and interacts with additional cognitive variables (e.g., *Gf*).

In addition, it would also be of interest to examine the possible influences of episodic memory retrieval on some of the effects we have documented. Recent studies have shown that an episodic specificity induction that biases reliance on episodic retrieval (Schacter & Madore, 2016) produces an increase in subsequent fluency and flexibility on the AUT (Madore et al., 2015; Madore et al., 2019). However, it is unknown whether or how inducing a reliance on episodic retrieval during the AUT would impact the semantic set-size effect or individual differences observed here. One possibility is that an episodic retrieval orientation mitigates the semantic set-size effect for SSC cues by providing access to episodic information in place of sparse semantic knowledge. Such studies could help to further our understanding of how both semantic and episodic memory contribute to creative cognition.

A few limitations exist in our study. First, current theories on the structure of semantic memory consider it dynamic, contingent on context and individual differences (Yee & Thompson-Schill, 2016); here, we maintain a more 'static' view of semantic memory structure. However, studies examining how properties of semantic memory network structure relate to individual differences in creative ability have shown consistent (i.e., static) characteristics in these networks, namely higher connectivity and shorter distances between concepts (Benedek et al., 2017; He et al., 2021; Ovando-Tellez et al., 2022). Furthermore, cognitive network research is slowly moving towards studying the dynamic nature of semantic memory structure (Bieth et al., 2021; Kenett & Thompson-Schill, 2020). A second limitation is that we used cues only from the University of South Florida Free Association Norms (Nelson et al., 2004). While future studies are needed to replicate our findings using other associative norms, we did replicate our findings across three studies and generalize them with an additional list. Third, although we assumed that RSC index semantic richness, our study could not directly isolate the relative roles of association number vs. dominance vs. structure on creative performance. These three aspects of semantic networks are related but nonredundant, so future work should future clarify their relative contribution to creative thinking. Fourth, we found relatively low inter-rater agreement in subjective creativity scoring in Study 2. Past work has reported similarly low agreement with subjective scoring, highlighting the need to employ objective assessments that can quantify originality with higher reliability, such as semantic distance (Beaty & Johnson, 2021). Fifth, participants in the in-lab study (Study 4) were considerably younger than participants in the online studies (Studies 1-3). Although *Gf* generally declines with age, we did not consider how age may interact with *Gf* or AUT performance, but we encourage future work to do so. Sixth, other potential individual differences beyond fluid intelligence may interact with the set-size findings presented here, such as openness to experience (Christensen et al., 2018). Future work should continue to examine individual differences to further uncover how memory systems and cognitive control interact to support creative thought. Finally, our classification of cue words into SSC and RSC was solely based on the number of associative responses, and we did not take into consideration aspects of forward and backward associative strength, information that is available in the University of South Florida Free Association Norms. Our decision to be agnostic to these different types of cue-response relationships may washes out specific asymmetric effects highlighted by Mak et al. (Mak et al., 2021; Mak & Twitchell, 2020). Future studies are needed to examine how the effects of set-size on AUT we find here are mediated by a cue word in- and out-degree centrality (forward and backward associative strength).

In conclusion, our research highlights a more nuanced and complex role of semantic memory in creative thinking. On the one hand, semantic memory is undoubtedly a critical infrastructure in generating creative ideas (Abraham & Bubic, 2015). On the other hand, however, similar to other cognitive processes that take part in creative thinking—such as cognitive control and working memory capacity—semantic memory can also constrain creative thinking. As such, our findings advance understanding of the complex orchestra of cognitive processes that give rise to creative thinking.

Acknowledgements

Roger E. Beaty is supported by a grant from the National Science Foundation [DRL-1920653]. Daniel L. Schacter is supported by the National Institute on Aging grant AG008441.

Conflict of Interest

The authors declare no conflict of interest.

Data availability

Data collected in these Studies are available upon request from the corresponding author.

References

Abraham, A. (2014). Creative thinking as orchestrated by semantic processing versus cognitive control brain networks [Perspective]. *Frontiers in Human Neuroscience*, *8*, 95. https://doi.org/10.3389/fnhum.2014.00095

Abraham, A., & Bubic, A. (2015). Semantic memory as the root of imagination. *Frontiers in Psychology*, *6*, 325. https://doi.org/10.3389/fpsyg.2015.00325

Acar, S., Chen, X., & Cayirdag, N. (2018). Schizophrenia and creativity: A meta-analytic review. *Schizophrenia Research*, *195*, 23-31. https://doi.org/https://doi.org/10.1016/j.schres.2017.08.036

Acar, S., & Runco, M. A. (2019). Divergent thinking: New methods, recent research, and extended theory. *Psychology of Aesthetics, Creativity and the Arts*, *13*(2), 153-158. https://doi.org/10.1037/aca0000231

Al-Azary, H., & Buchanan, L. (2017). Novel metaphor comprehension: Semantic neighbourhood density interacts with concreteness [journal article]. *Memory & Cognition*, *45*(2), 296-307. https://doi.org/10.3758/s13421-016-0650-7

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*(4), 451-474.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186-197.

Avitia, M. J., & Kaufman, J. C. (2014). Beyond *g* and *c*: The relationship of rated creativity to long-term storage and retrieval (*Glr*). *Psychology of Aesthetics, Creativity and the Arts*, *8*(3), 293-302. https://doi.org/10.1037/a0036772

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.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316.

Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). Reasoned connections: A dual-process perspective on creative thought. *Thinking & Reasoning*, *21*(1), 61-75.

Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences*, *20*(2), 87-95. https://doi.org/10.1016/j.tics.2015.10.004

Beaty, R. E., Christensen, A. P., Benedek, M., Silvia, P. J., & Schacter, D. L. (2017). Creative constraints: Brain activity and network dynamics underlying semantic interference during idea production. *NeuroImage*, *148*, 189-196. https://doi.org/10.1016/j.neuroimage.2017.01.012

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

Beaty, R. E., & Silvia, P. J. (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. https://doi.org/https://psycnet.apa.org/doi/10.1037/a0029171

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition*, *42*(7), 1-12. https://doi.org/10.3758/s13421-014-0428-8

Beaty, R. E., Zeitlen, D. C., Baker, B. S., & Kenett, Y. N. (2021). Forward flow and creative thought: Assessing associative cognition and its role in divergent thinking. *Thinking Skills and Creativity*, *41*, 100859. https://doi.org/https://doi.org/10.1016/j.tsc.2021.100859

Benedek, M., & Fink, A. (2019). Toward a neurocognitive framework of creative cognition: the role of memory, attention, and cognitive control. *Current Opinion in Behavioral Sciences*, *27*, I116-122. https://doi.org/https://doi.org/10.1016/j.cobeha.2018.11.002

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73-83. https://doi.org/http://dx.doi.org/10.1016/j.intell.2014.05.007

Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2017). How semantic memory structure and intelligence contribute to creative thought: a network science approach. *Thinking & Reasoning*, *23*(2), 158-183. https://doi.org/10.1080/13546783.2016.1278034

Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's model on creativity-related differences in associative hierarchies. Evidence for a common path to uncommon thought. *The Journal of Creative Behavior*, *47*(4), 273-289. https://doi.org/10.1002/jocb.35

Bieth, T., Kenett, Y. N., Ovando-Tellez, M., Lopez-Persem, A., Lacaux, C., Oudiette, D., & Volle, E. (2021). Dynamic changes in semantic memory structure support successful problem-solving. *PsyArXiv*. https://doi.org/https://doi.org/10.31234/osf.io/38b4w

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3-5. https://doi.org/10.1177/1745691610393980

Bunting, M. F., Conway, A. R. A., & Heitz, R. P. (2004). Individual differences in the fan effect and working memory capacity. *Journal of Memory and Language*, *51*(4), 604-622.

Christensen, A. P., Kenett, Y. N., Cotter, K. N., Beaty, R. E., & Silvia, P. J. (2018). Remotely close associations: Openness to experience and semantic memory structure. *European Journal of Personality*, *32*(4), 480-492. https://doi.org/doi:10.1002/per.2157

Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology*, *53*(2), 82-88.

Chrysikou, E. G. (2018). The costs and benefits of cognitive control for creativity. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge Handbook of the Neuroscience of Creativity* (pp. 299-317). Cambridge University Press.

Chrysikou, E. G. (2019). Creativity in and out of (cognitive) control. *Current Opinion in Behavioral Sciences*, *27*, 94-99. https://doi.org/https://doi.org/10.1016/j.cobeha.2018.09.014

Chrysikou, E. G., Motyka, K., Nigro, C., Yang, S.-I., & Thompson-Schill, S. L. (2016). Functional fixedness in creative thinking tasks depends on stimulus modality. *Psychology of Aesthetics, Creativity and the Arts*, *10*(4), 425-435. https://doi.org/10.1037/aca0000050

Chrysikou, E. G., & Weisberg, R. W. (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. https://doi.org/10.1037/0278-7393.31.5.1134

Cofer, C. N., & Shevitz, R. (1952). Word-associations as a function of word frequency. *The American Journal of Psychology*, *62*(1), 75-79.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.

Danguecan, A. N., & Buchanan, L. (2016). Semantic neighborhood effects for abstract versus concrete words. *Frontiers in Psychology*, *7*, 1034.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words [journal article]. *Behavior Research Methods*, *51*(3), 987-1006. https://doi.org/10.3758/s13428-018-1115-7

Dumas, D., Organisciak, P., & Doherty, M. (2021). Measuring originality with human raters and text-mining models: A psychometric comparison of methods. *Psychology of Aesthetics, Creativity and the Arts*, *15*(4), 645-663. https://doi.org/https://doi.org/10.1037/aca0000319

Duñabeitia, J. A., Avilés, A., & Carreiras, M. (2008). NoA's ark: Influence of the number of associates in visual word recognition. *Psychonomic Bulletin & Review*, *15*(6), 1072-1077.

Faust, M., & Kenett, Y. N. (2014). Rigidity, chaos and integration: Hemispheric interaction and individual differences in metaphor comprehension [Hypothesis & Theory]. *Frontiers in Human Neuroscience*, *8*(511), 1-10. https://doi.org/10.3389/fnhum.2014.00511

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. https://doi.org/http://dx.doi.org/10.1016/j.intell.2016.03.005

Gentner, D. (1981). Verb semantic structures in memory for sentences: Evidence for componential representation. *Cognitive Psychology*, *13*(1), 56-83.

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. https://doi.org/10.1348/096317907X173421

Glucksberg, S., & Weisberg, R. W. (1966). Verbal behavior and problem solving: Some effects of labeling in a functional fixedness problem. *Journal of Experimental Psychology*, *71*(5), 659-664.

Gray, K., Anderson, S., Chen, E. E., Kelly, J. M., Christian, M. S., Patrick, J., Huang, L., Kenett, Y. N., & Lewis, K. (2019). "Forward flow": A new measure to quantify free thought and predict creativity. *American Psychologist*, *74*(5), 539-554. https://doi.org/https://psycnet.apa.org/doi/10.1037/amp0000391

Green, A. E. (2016). Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. *Current Directions in Psychological Science*, *25*(1), 28-35. https://doi.org/10.1177/0963721415618485

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006-1033.

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., Halpern, D., Hamrick, J. B., & Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, *48*(3), 829-842.

Hass, R. W. (2017). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory & Cognition*, *45*(2), 233-244. https://doi.org/10.3758/s13421-016-0659-y

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

Hass, R. W., Rivera, M., & Silvia, P. J. (2018). On the dependability and feasibility of layperson ratings of divergent thinking [Original Research]. *Frontiers in Psychology*, *9*(1343). https://doi.org/10.3389/fpsyg.2018.01343

He, L., Kenett, Y. N., Zhuang, K., Liu, C., Zeng, R., Yan, T., Huo, T., & Qiu, J. (2021). The relation between semantic memory structure, associative abilities, and verbal and figural creativity. *Thinking & Reasoning*, *27*(2), 268-293. https://doi.org/10.1080/13546783.2020.1819415

Heinen, D. J. P., & Johnson, D. R. (2018). Semantic distance: An automated measure of creativity that is novel and appropriate. *Psychology of Aesthetics, Creativity and the Arts*, *12*(2), 144-156. https://doi.org/10.1037/aca0000125

Hills, T. T., & Kenett, Y. N. (2022). Is the Mind a Network? Maps, Vehicles, and Skyhooks in Cognitive Network Science. *Topics in Cognitive Science*, *14*(1), 189-208. https://doi.org/https://doi.org/10.1111/tops.12570

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66-71.

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment* (Vol. 53). John Wiley & Sons.

Kenett, Y. N. (2018a). Going the extra creative mile: The role of semantic distance in creativity – theory, research, and measurement. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge Handbook of the Neuroscience of Creativity* (pp. 233-248). Cambridge University Press.

Kenett, Y. N. (2018b). Investigating creativity from a semantic network perspective. In Z. Kapoula, E. Volle, J. Renoult, & M. Andreatta (Eds.), *Exploring Transdisciplinarity in Art and Sciences* (pp. 49-75). Springer International Publishing. https://doi.org/10.1007/978-3-319-76054-4_3

Kenett, Y. N. (2019). What can quantitative measures of semantic distance tell us about creativity? *Current Opinion in Behavioral Sciences*, *27*, 11-16. https://doi.org/https://doi.org/10.1016/j.cobeha.2018.08.010

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, *8*(407), 1-16. https://doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., & Austerweil, J. L. (2016). *Examining search processes in low and high creative individuals with random walks* Proceedings of the 38th Annual Meeting of the Cognitive Science Society, Austin, TX.

Kenett, Y. N., Beaty, R. E., Silvia, P. J., Anaki, D., & Faust, M. (2016). Structure and flexibility: Investigating the relation between the structure of the mental lexicon, fluid intelligence, and creative achievement. *Psychology of Aesthetics, Creativity and the Arts*, *10*(4), 377-388. https://doi.org/10.1037/aca0000056

Kenett, Y. N., & Faust, M. (2019). A semantic network cartography of the creative mind. *Trends in Cognitive Sciences*, *23*(4), 271-274. https://doi.org/10.1016/j.tics.2019.01.007

Kenett, Y. N., Levy, O., Kenett, D. Y., Stanley, H. E., Faust, M., & Havlin, S. (2018). Flexibility of thought in high creative individuals represented by percolation analysis. *Proceedings of the National Academy of Sciences*, *115*(5), 867-872. https://doi.org/10.1073/pnas.1717362115

Kenett, Y. N., & Thompson-Schill, S. L. (2020). Novel conceptual combination can dynamically reconfigure semantic memory networks. *PsyArXiv*. https://doi.org/10.31234/osf.io/crp47

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

Klimesch, W. (1987). A connectivity model for semantic processing. *Psychological Research*, *49*(1), 53-61. https://doi.org/10.1007/BF00309203

Kroll, N. E. A., & Klimesch, W. (1992). Semantic memory: Complexity or connectivity? *Memory & Cognition*, *20*(2), 192-210.

Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic Bulletun & Review*, *28*(1), 40-80. https://doi.org/10.3758/s13423-020-01792-x

Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology*, *8*(3), 311-335.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1-36. https://journals.lww.com/ear-hearing/Fulltext/1998/02000/Recognizing_Spoken_Words__The_Neighborhood.1.aspx

Madore, K. P., Addis, D. R., & Schacter, D. L. (2015). Creativity and memory: Effects of an episodic-specificity induction on divergent thinking. *Psychological Science*, *26*(9), 1461-1468. https://doi.org/10.1177/0956797615591863

Madore, K. P., Thakral, P. P., Beaty, R. E., Addis, D. R., & Schacter, D. L. (2019). Neural mechanisms of episodic retrieval support divergent creative thinking. *Cerebral Cortex*, *29*(1), 150-166.

Mak, M. H. C., Hsiao, Y., & Nation, K. (2021). Lexical connectivity effects in immediate serial recall of words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(12), 1971-1997. https://doi.org/10.1037/xlm0001089

Mak, M. H. C., & Twitchell, H. (2020). Evidence for preferential attachment: Words that are more well connected in semantic networks are better at acquiring new links in paired-associate learning. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-020-01773-0

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57-78. https://doi.org/http://dx.doi.org/10.1016/j.jml.2016.04.001

Marko, M., & Riečanský, I. (2021). The structure of semantic representation shapes controlled semantic retrieval. *Memory*, *29*(4), 538-546.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1996). How examples may (and may not) constrain creativity. *Memory & Cognition*, *24*(5), 669-680.

Matheson, H. E., & Kenett, Y. N. (2021). A novel coding scheme for assessing responses in divergent thinking: An embodied approach. *Psychology of Aesthetics, Creativity and the Arts*, *15*(3), 412-425. https://doi.org/10.1037/aca0000297

Mednick, S., A. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220-232. http://www.ncbi.nlm.nih.gov/pubmed/14472013

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65-79.

Nelson, D. L., & McEvoy, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 292-314. https://doi.org/https://psycnet.apa.org/doi/10.1037/0278-7393.5.3.292

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407. https://doi.org/10.3758/BF03195588

Nelson, D. L., Schreiber, T. A., & Xu, J. (1999). Cue set size effects: Sampling activated associates or cross-target interference? *Memory & Cognition*, *27*(3), 465-477.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, *39*(1), 36-45.

Ovando-Tellez, M., Kenett Yoed, N., Benedek, M., Bernard, M., Belo, J., Beranger, B., Bieth, T., & Volle, E. (2022). Brain connectivity–based prediction of real-life creativity is mediated by semantic memory structure. *Science Advances*, *8*(5), eabl4294. https://doi.org/10.1126/sciadv.abl4294

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203. https://doi.org/10.3758/s13428-018-01193-y

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*(1), 161-167.

Prabhakaran, R., Green, A. E., & Gray, J. R. (2014). Thin slices of creativity: Using single-word utterances to assess creative cognition. *Behavior Research Methods*, *46*(3), 641-659. https://doi.org/10.3758/s13428-013-0401-7

Radvansky, G. A. (1999). Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General*, *128*(4), 563-579.

Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts [Original Research]. *Frontiers in Human Neuroscience*, *6*(315), 315. https://doi.org/10.3389/fnhum.2012.00315

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2007). Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas. *Journal of Experimental Social Psychology*, *43*(6), 933-946. https://doi.org/https://doi.org/10.1016/j.jesp.2006.10.014

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, *24*(1), 66-75. https://doi.org/10.1080/10400419.2012.652929

Schacter, D. L., & Madore, K. P. (2016). Remembering the past and imagining the future: Identifying and enhancing the contribution of episodic memory. *Memory Studies*, *9*(3), 245-255.

Schilling, M. A. (2005). A "small-world" network model of cognitive insight. *Creativity Research Journal*, *17*(2-3), 131-154. https://doi.org/10.1080/10400419.2005.9651475

Siew, C. S. Q. (2013). Community structure in the phonological network [Original Research]. *Frontiers in Psychology*, *4*, 553. https://doi.org/10.3389/fpsyg.2013.00553

Siew, C. S. Q. (2019). spreadr: An R package to simulate spreading activation in a network. *Behavior Research Methods*, *51*(2), 910-929.

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence*, *41*(5), 328-340. https://doi.org/10.1016/j.intell.2013.05.004

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American Journal of Psychology*, *104*(1), 61-87.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, *21*(6), 837-845.

Sowden, P. T., Pringle, A., & Gabora, L. (2014). The shifting sands of creative thinking: Connections to dual-process theory. *Thinking & Reasoning*, *21*(1), 40-60. https://doi.org/10.1080/13546783.2014.885464

Stella, M., & Kenett, Y. N. (2019). Viability in multiplex lexical networks and machine learning characterizes human creativity. *Big Data and Cognitive Computing*, *3*(3), 45. https://doi.org/10.3390/bdcc3030045

Torrance, E. P. (1972). Predictive validity of the Torrance tests of creative thinking. *The Journal of Creative Behavior*, *6*(4), 236-262.

Van Stockum, C. A., & DeCaro, M. S. (2020). When working memory mechanisms compete: Predicting cognitive flexibility versus mental set. *Cognition*, *201*, 104313. https://doi.org/https://doi.org/10.1016/j.cognition.2020.104313

Volle, E. (2018). Associative and controlled cognition in divergent thinking: Theoretical, experimental, neuroimaging evidence, and new directions. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge handbook of the neuroscience of creativity* (pp. 333-362). Cambridge University Press.

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

Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition*, *26*(4), 716-730. https://doi.org/10.3758/BF03211392

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. (2012). An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 72.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*(4), 1015-1027. https://doi.org/10.3758/s13423-015-0948-7