Thinking & Reasoning

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ptar20

Determinants of creative thinking: the effect of task characteristics in solving remote associate test problems

Ut Na Sio, Kenneth Kotovsky & Jonathan Cagan

To cite this article: Ut Na Sio, Kenneth Kotovsky & Jonathan Cagan (2021): Determinants of creative thinking: the effect of task characteristics in solving remote associate test problems, Thinking & Reasoning, DOI: 10.1080/13546783.2021.1959400

To link to this article: https://doi.org/10.1080/13546783.2021.1959400

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 09 Aug 2021.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=ptar20

THINKING & REASONING
https://doi.org/10.1080/13546783.2021.1959400

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Determinants of creative thinking: the effect of task characteristics in solving remote associate test problems

Ut Na Sio[a], Kenneth Kotovsky[b] and Jonathan Cagan[c]

[a]Sheffield University Management School, The University of Sheffield, Sheffield, UK; [b]Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; [c]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

ABSTRACT

Creative problem solving is often conceptualised as a process of search. However, little is known about the difficulties of carrying out this search process. We conducted three studies examining how strongly different task characteristics impact creative problem-solving performance. In Study 1, regression analyses on normative data of Remote Associates Test (RAT) problems identified key task characteristics that best predict performance. We replicated these findings in a sample that was more diverse with respect to age and education background and proposed that two key factors may interact in predicting RAT problem difficulty (Study 2). We then confirmed this prediction in a pre-registered study (Study 3). Our results suggest that (a) the semantic distance between the cues and the answer and (b) the number of strong but irrelevant associates are important determinants of RAT problem difficulty, and that their influence is interdependent. Implications for theories and for studies aimed at improving creative problem-solving performance are discussed.

ARTICLE HISTORY Received April 7, 2020; Accepted July 17, 2021

KEYWORDS Creative thinking; problem solving; task characteristics; remote associates test

Creative problem solving—the ability to generate novel and feasible solutions—is central to many scientific breakthroughs. It is also considered one of the most important skills needed in the future workforce (World

CONTACT Ut Na Sio u.n.sio@sheffield.ac.uk Sheffield University Management School, The University of Sheffield, Conduit Road, Sheffield, S10 1FL, UK.

All materials and data are available through the Open Science Framework: https://osf.io/6cwgv.

Supplemental data for this article is available online at https://doi.org/10.1080/13546783.2021.1959400.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Economic Forum, 2018). Unfortunately, thinking creatively is difficult. Individuals, even experts, often perform sub-optimally when solving creative problems (Bilalić et al., 2008). Finding ways to facilitate the solution search processes involved in creative problem solving has been a key question in many domains, such as cognition, education, engineering design and management science. However, little is known about the factors that impact these solution search processes. Exploring this question is important as it could inform future studies aimed at enhancing creative problem-solving performance. We present three studies aiming at quantifying the influence of different task characteristics on creative problem-solving performance.

Task characteristics of creative problem

Creative problem solving, similar to other types of problem solving, can be conceptualised as a process of search. During problem solving, individuals will first encode the problem and generate an initial problem space consisting of concepts related to the problem, e.g., prior knowledge on solving similar problems. Individuals then search for a solution within this space (Newell & Simon, 1972). When solving routine or well-defined problems, prior knowledge can help to confine the search space and guide individuals in the right direction (Bédard & Chi, 1975). However, many problems encountered in real life often require novel solutions, requiring individuals to search beyond the typical solutions (Beaty et al., 2019; Beaty & Silvia, 2012; Gilhooly et al., 2007; Goel & Pirolli, 1992; Paton & Dorst, 2011; Smith et al., 2013). Exploring the problem space effectively is thus critical for creative problem-solving. An increasing number of studies have been conducted to better understand the search processes involved in successful creative problem-solving.

It has been suggested that creative thinking requires two distinct search processes—the divergent and convergent search processes (e.g., Brophy, 1998). The divergent one can be viewed as the process of generating many different ideas and the convergent one can be viewed as the process of generating one possible solution that satisfies a number of criteria (Guildford, 1967; Hommel, 2012). The present work focussed on the convergent search process by examining how different task characteristics impact the performance on Remote Associates Test problem (RAT, Mednick, 1962), a common task used to assess creative convergent thinking (Brophy, 1998; Hommel, 2012; Lee & Therriault, 2013). The findings should allow us to identify the key difficulties associated with the convergent creative thinking process and in turn can help determine the convergent search process critical to successful solution discovery.

RAT problem solving

RAT problems require finding a word that can be associated with the three given cues (Mednick, 1962). When solving RAT problems, individuals have to search within a problem space of words associated to the given cues to look for a promising answer or expand their search if the solution is outside this initial problem space. RAT problems have been designed to assess individuals' ability to retrieve remote associates, which is considered to be critical in creative thinking (Mednick, 1962). Positive correlations have been reported between performance on RAT problem-solving and other creative problem-solving tasks, such as associative fluency (Levin, 1978; Mednick et al., 1964), rebus (MacGregor & Cunningham, 2008), verbal insight problem and analogical transfer (Ansburg & Dominowski, 2000; Cushen & Wiley, 2018). RAT problem solving is also viewed as a measure of convergent creative thinking because it involves finding a single correct solution to a problem where there are usually multiple apparent initial pathways (Brophy, 1998; Hommel, 2012; Lee & Therriault, 2013). At its core, RAT problem solving taps basic semantic retrieval processes, a cognitive operation that is central to many types of problem solving (Davelaar, 2015; Kenett et al., 2017; Smith et al., 2013; Smith & Vul, 2015). RAT problems have also been used to evaluate interventions intended to improve creative problem solving, e.g., sleep, incubation, task-alternation, and mind-wandering (Cai et al., 2009; Lu et al., 2017; Sio & Ormerod, 2009; Threadgold et al., 2019; Zedelius & Schooler, 2015; see Wu et al., 2020, for a review).

Because of the nature of RAT problem solving, examining the influence of various task characteristics on RAT problem-solving performance should improve our understanding of difficulties associated with the convergent search process in creative problem solving and in other problem-solving domains where associative processing is critical. The findings of the present work should also enable a better understanding of the mechanisms of different creative problem-solving interventions, given that RAT problems are commonly used to evaluate their effectiveness.

Although numerous studies have been conducted to examine how task characteristics could impact RAT problem-solving performance, the majority of them focussed on the associative properties of the cues and the answer, e.g., the semantic distance between the cue and the answer, the frequency of the answer, and the number of irrelevant ideas (Gupta et al., 2012; Olteīeanu & Schultheis, 2017; Sio et al., 2013; Wiley, 1998; Wu et al., 2017). Only recently studies have started to examine the link between semantic network structure and creative problem-solving performance (Kenett et al., 2014, 2018), despite research that has long suggested its significant role in the creative process (Abraham & Bubic, 2015; Benedek et al., 2012;

Mednick, 1962; Mendelsohn, 1976, Schilling, 2005; for a discussion, see Kenett, 2018). However, these studies were primarily concerned with individual differences in semantic network structure and creative problem-solving ability. As such little is known about how the structure of the initial problem space impacts the search process. In addition, past studies have mainly focussed on examining the impact of one single characteristic on RAT problem solving, an approach that may not provide an accurate description of RAT problem solving if different task characteristics are inter-related or interdependent in influencing RAT problem difficulty. To address these issues, we conducted three studies to investigate the influence of multiple characteristics on RAT problem solving, allowing us to better estimate their individual impact. Both characteristics related to the associative properties of the cues and the answer as well as to the structure of the initial problem space were considered. The characteristics studied are:

Typicality of the answer

During problem solving, individuals will first generate a problem space that consists of typical concepts that are closely related to the problem and have a high frequency of occurrence (Gupta et al., 2012; Landau & Lehr, 2004; Wiley, 1998). When the correct answer for the problem is a typical idea, successful solution requires only a narrow search within the initial problem space. In contrast, when a problem's answer is less typical, individuals need to expand the search space to reach the correct answer. As such, the typicality of the answer is likely to predict problem difficulty (Davelaar, 2015; Gruszka & NecKa, 2002; Marko et al., 2019; Sio et al., 2013).

Number of typical but irrelevant associates

As mentioned above, during problem solving, individuals often first retrieve concepts that are closely related to the problem. For problems requiring novel solutions, retrieving these close associates may not always lead to an effective outcome; instead, it may prevent the retrieval of the solution (Luchins & Luchins, 1959). Inhibition of these strong but irrelevant associates is often critical for successful solution discovery, particularly when these irrelevant associates are prevalent (Benedek et al., 2014; Bubic et al., 2010; Jansson & Smith, 1991; Moss et al., 2011; Smith & Blankenship, 1991; Storm et al., 2011; Unsworth et al., 2011). Due to this additional inhibition required, we expected that problems with a high number of irrelevant associated should be more difficult to solve.

Structure of the problem space

If problem solving is a process of search within a problem space, problem solving performance may also depend on the structure of the space

(Abraham & Bubic, 2015; Benedek et al., 2012; Mednick, 1962; Mendelsohn, 1976; Schilling, 2005). Two common measures of the structure of a problem space are the level of clustering and the path length between associates (Humphries & Gurney, 2008). A problem space with a high level of clustering implies that a shift from one cluster to another would be needed to exhaust all the associates in the space. Past studies have shown that individuals tended not to change the search direction once an investment in effort or time has been made (Viswanathan & Linsey, 2013). Given the reluctance to switch search direction, there may be a negative link between the level of clustering and RAT problem difficulty. Another structural characteristic that we examined was the path length between associates. Having long path lengths between associates implies that many associates are not directly connected but linked indirectly through intermediate associates (Latora & Marchiori, 2001). For such problems, it may take longer time to explore the search space thoroughly.

We conducted three studies to evaluate which of these characteristics have the most influence on RAT problem-solving performance with the aim of improving the current understanding of the difficulties associated with the solution search process involved in creative problems. Note that RAT problems may also vary in other ways, but we focussed on those that are theoretically critical to the search process in RAT problem solving.

Study 1

Study 1 explored which of the task characteristics mentioned above are the most influential determinants of RAT problem solving difficulty.

Method

Task

In this study, RAT problems that require finding a word that forms a compound word or phrase with the given cues (e.g., cues: stick, maker, point; answer: match) were used as the problem-solving tasks. To explore which task characteristics best predict performance on RAT problem solving, we analysed the normative dataset created by Bowden and Jung-Beeman (2003b), which provides the mean solution time and solution rate for 144 compound RAT problems, with four different time limits (2 sec, 7 sec, 15 sec, 30 sec). Our analysis focussed on the solution rate in the 30-sec condition as this condition provided the greatest amount of time for individuals to conduct the solution search.

Task characteristics

For each RAT problem, a problem space using the R package "igraph" (Csardi & Nepusz, 2006) was constructed. The problem space included the three given cues, the answer, and associates of the three cues, as reported in the free association norms (Nelson et al., 1998). Words in the space were inter-connected according to the association among them: the connection weight from word *a* to word *b* ($w_{a \rightarrow b}$) was equal to the strength of the association from word *a* to word *b* listed in the free association norms (Nelson et al., 1998). Note that the free association norms are likely to omit very weak associates of the cues. However, this limitation is not critical for this study as we focussed the analysis on the characteristics of the *initial* problem space (i.e., generated at the beginning of problem solving) and it has been suggested that this initial problem space includes mostly close associates (Newell & Simon, 1972). We believe the associative spaces based on the free-association norms provide good representations of the initial problem spaces.

Twenty-seven of the 144 RAT problems from the normative dataset had cues words that are not listed in the free association norms and were excluded as it would be impossible to compute their task characteristics accurately. Two additional problems were excluded due to a lack of performance data in the 30 sec time condition. Our final problem set included 115 RAT problems. For each problem, three characteristics— (1) the typicality of the answer, (2) the number of typical but irrelevant associates in the initial problem space, and (3), the structure of the initial problem space— were examined. Each characteristic and the various ways they were operationalised are described below.

1. **Typicality of the Answer.** For each RAT problem, the typicality of the answer was measured in terms of:
    1.1. ***Frequency of the Answer Words.*** This was assessed in three ways. One was to count the *written word frequency* (based on the Corpus of Contemporary American English, COCA). The second measure was the *associate frequency* (AF; Griffiths et al., 2007; Gupta et al., 2012), which was the sum of the associative strengths of all the words that are associated with the answer word. The associative strengths were based on the free associative norms (Nelson et al., 1998). The third measure was *the average frequency of the cue-answer compound words/phrases,* which was determined by averaging the frequency of occurrence of each three cue-answer compound words/phrases in the COCA corpus. Past studies have used this metric as a measure of the typicality of the RAT problem answers (Olteţeanu & Falomir, 2015; Olteīeanu & Schultheis, 2017).

1.2. ***Semantic Distance Between the Cues and the Answer.*** This was measured in two ways. One was to count the number of cues directly associated to the answer, as determined by the free associative norms (i.e., *number of cue-answer pairs)*. The other was to sum the associative strengths from the cues to the answer (i.e., *strength of cue-answer association)*. These two measures are commonly used to assess how closely related the cues and the answer are in the RAT problems (Sio et al., 2013; Wu et al., 2017).

2. **Number of Typical but Irrelevant Associates.** Two measures were used to count the number of these associates:

   2.1. ***Number of Strong but Irrelevant Associates.*** When solving RAT problems, individuals often have to disregard associates that are strongly connected with the cues but irrelevant to the answer (Moss et al., 2011; Smith & Blankenship, 1991; Storm et al., 2011). To compute the number of these strong but irrelevant associates, we counted, for each RAT problem, the number of words whose association with the cue was stronger than the association between the cue and the answer.

   2.2. ***Number of Misleading Associates***. For each RAT problem, misleading associates were operationalised as concepts that are closely related to only two out of the three cues. For example, one misleading associate for the RAT problem "stick, point, maker", is "pin", which is associated with "stick" and "point" but not "maker". It has been suggested that when solving RAT problems, individuals may fixate on these misleading associates, blocking the retrieval of the correct answer (Sio et al., 2017; Wiley, 1998).

3. **Structure of the Problem Space.** For each RAT problem, the structure of the problem space was assessed in terms of:

   3.1. ***Level of Clustering.*** To measure the extent to which associates in the problem space tend to cluster together, a clustering coefficient was computed using the R package "tnet"—a package for computing the clustering coefficient of a weighted network (Opsahl & Panzarasa, 2009). To compute the coefficient, the number of triplets—three associates that are connected by either two direct links (i.e., forming an open triangle) or three direct links (i.e., forming a closed triangle)—was first counted. Each triplet was then weighted based on the associative strengths between the connected associates. A clustering coefficient was then computed by dividing the number of weighted closed triplets by the number of weighted open and closed triplets. These clustering coefficients are likely positively correlated with the size of the problem space. To correct for this size bias, *a relative clustering coefficient* was computed by dividing the coefficient by the coefficient of an equivalent random network—a network with the same number of nodes and edges, and all associative strengths equal to 1 (Humphries & Gurney, 2008). This relative clustering coefficient was used as the measure of the level of clustering of the problem space in this study.

3.2. ***Average Shortest Path Length (ASP) between Associates of the Space***. Path length represents the number of steps needed to traverse from one node to another node in the network. In our study, for any two associates in the problem space, e.g., A and B, if A is directly associated to B with an associative strength equal to $w_{A \to B}$, the shortest path length between them would be 1- $w_{A \to B}$. This is to capture the intuition that semantically-close associates would have short path length between them (Kenett et al., 2017). If the shortest path between A and B is a path through the intermediate node, e.g., A→C→B, the shortest path length between them would be equal to the sum of the path length from A→C and C→B. If there was no such path between two associates, the shortest path length between them would be one longer than the longest of all the shortest paths in the space (Csardi & Nepusz, 2006). A shorter ASP indicates a more well-connected network. The ASP was computed using the R package "igraph", a package for analysing networks (Csardi & Nepusz, 2006). The ASP was then calculated by taking the arithmetic average of these shortest path lengths. Again, to correct for the size bias, a relative ASP was computed by dividing the ASP of the space with the ASP of an equivalent random network.

Results and discussion

A backward stepwise regression analysis was conducted to identify which task characteristics significantly impact RAT problem-solving performance. The final, most optimal, model was determined using the Akaki information Criterion (AIC) as the selection criteria. The dependent variable was the 30-second solution rate of the RAT problems reported in the normative set (Bowden & Jung-Beeman, 2003b). The predictor variables were the nine measures of task characteristics described above. The *M* and the *SD* of each measure of task characteristics and the correlations between them are shown in Table 1. Task characteristic measures were rescaled (i.e., rescaled to $M = 0$, $SD = 1$). Table 2 presents the initial (including all the predictor variables) and final models. For the final model, the small variance inflation factors (VIFs), ranging from 1.7 to 1.9, suggest that multicollinearity was not a concern. Assumptions of normality and homoscedasticity of model residuals were respected (Breusch-Pagan test for heteroscedasticity: $p = .37$, Anderson-Darling test for normality: $p = .19$).

Three predictor variables—the number of cue-answer pairs, the strength of cue-answer pairs, and the number of strong but irrelevant associates—remained in the final optimal model. The final model accounted for 33% of variance, $F(3,111) = 18.02$, $p < .001$. The strength of the cue-answer pairs and the number of cue-answer pairs were both positively associated with RAT problem-solving performance but the effect of the number of the cue-answer pairs was not statistically significant (strength of cue-answer pairs: $\beta = .30$, $p < .01$; number of cue-answer pairs: $\beta = .17$, $p = .11$). The seemingly positive relations between these variables and RAT problem-solving performance are consistent with past findings that the semantic distance between the answer and the cues is a key determinant of RAT problem difficulty (Davelaar, 2015; Gruszka & NecKa, 2002; Marko et al., 2019; Sio et al., 2013). The presence of strong but irrelevant associates of the problem is viewed as another source of difficulties in RAT problem solving (Moss et al., 2011; Smith & Blankenship, 1991; Storm et al., 2011). In line with the literature, the number strong but irrelevant associates remained in the final model. It was found to be negatively related to RAT performance, but the effect was only marginally significant, $\beta = -.20$, $p = .06$. As a test of robustness, a similar regression analysis, but with a forward selection method was also conducted and the analysis selected the same key characteristics.

Study 1 has identified that the semantic distance between the problem and the answer, and the number of strong but irrelevant associates were the key characteristics that determine RAT problem difficulty. One may question the generalisability of these findings, given that the normative data used were collected from a relatively small sample of university students in controlled lab settings. To address this issue, Study 2 was conducted in an attempt to replicate the findings of Study 1 using a more representative sample of participants, tested in less controlled environment.

Study 2

Study 2 was conducted to determine if the factors identified in Study 1 would also be predictive of performance in a new and more diverse sample of participants.

Table 1. Descriptive statistics and correlations between the task characteristics measures and RAT problem solution rate (normative data).

Task Characteristics Measures	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Written word frequency of the answer word	1.15×10^5	1.46×10^5									
2. Associate frequency of the answer word	3.98	4.30	.25**								
3. Avg. frequency of the cue-answer compound words/phrases	1.07×10^3	1.41×10^3	.10	.08							
4. Strength of association from the cues to the answer	0.97	0.15	−.02	.16+	.02						
5. Number of cue-answer pairs	1.08	0.88	−.05	.08	.12	.57**					
6. Number of strong but irrelevant associates	36.16	13.01	−.01	-.07	-.06	-.60**	-.60**				
7. Number of misleading associates	0.63	0.80	.01	.06	-.02	-.02	-.03	.23*			
8. Relative clustering coefficient	6.78	1.78	.01	-.23*	-.07	-.28**	-.20*	.32**	-.05		
9. Relative ASP	1.29	0.23	.06	-.07	-.09	-.14	-.11	.19*	-.01	.44**	
10. Solution Rate	.45	.25	−.02	< .01	.09	.52**	.46**	-.48**	-.08	-.14	-.09

**$p < .01$.
*$p < .05$.
+$p < .10$.

Table 2. The initial and final regression models.

Predictors	Initial Model			Final Model		
	β	t	p	β	t	p
(Intercept)		3.21	$<$ .01		5.11	$<$ .001
Typicality of the Answer						
Frequency of the Answer Word						
Written word frequency of the answer word	$<$ .01	0.03	.97		Excluded	
Associate frequency of the answer word	-.08	−0.88	.38		Excluded	
Avg. frequency of the cue-answer compound words/phrases	.06	0.67	.50		Excluded	
Semantic Distance Between the Cues and the Answer						
Strength of cue-answer pairs	.33	3.04	$<$ .01	.30	2.94	$<$ .01
Number of cue-answer pairs	.16	1.48	.14	.17	1.62	.11
Number of Typical but Irrelevant Associates						
Number of strong but irrelevant associates	-.20	−1.70	.09	-.20	−1.88	.06
Number of misleading associates	-.01	−0.15	.88		Excluded	
Structure of the Problem Space						
Relative clustering coefficient	.03	0.33	.74		Excluded	
Relative ASP	-.01	−0.08	.94		Excluded	

Note. The initial model was significant, $F(9,105) = 5.93$, $p < .001$. adjusted $R^2 = .28$, AIC = -350.97. The final model was significant, $F(3,111) = 18.02$, $p < .001$. adjusted $R^2 = .31$, AIC = -361.32.

Method

Participants

A sample of 156 native English speakers (82 females, 74 males) residing in the United States was recruited via MTurk. The mean age of the participants was 39.00 years ($SD = 10.42$). All were MTurk workers with an approval rate (the percentage of previous HITs that were approved) of at least 95%. About half of the participants (48.1%) reported having a college degree or above. They were paid USD $2 for completing the study, which on average took 20 minutes.

Task

Sixty Remote Associate Test (RAT) problems were selected from the normative set used in Study 1 (Bowden & Jung-Beeman, 2003b). Supplementary Table S1 presents the selected RAT problems. The following measures of task characteristics, which were shown to be relevant to problem solving performance in Study 1, were examined:

Typicality of the answer. This was measured in terms of the *number of cue-answer pairs* and *the strength of cue-answer pairs.*

Number of typical but irrelevant associates. This was measured by counting *the number of strong but irrelevant associates*.

The 60 RAT problems, varied along these characteristics, were divided into 3 sets, which were matched on the task characteristics. See supplementary materials (Table S2) for the *M* and the *SD* of each characteristic measure for each set.

Procedures

Participants were randomly assigned to solve one of the three sets of 20 RAT problems. Each RAT problem was presented once for 30 s. During each problem, participants could enter their answers at any point. If their response was correct, the next problem was presented; otherwise, they had the remainder of the time limit to continue working on the problem. The RAT problems were presented in a randomised order. The study was delivered via Gorilla.sc (Anwyl-Irvine et al., 2020).

Results and discussion

The average solution rate for the RAT problems in this study was 0.43 ($SD = .17$). There was a moderate and positive correlation between our solution rates and the normative solution rates (Bowden & Jung-Beeman, 2003b), $r = 0.63$, $p < .001$, indicating that the relative difficulty of RAT problems was

Table 3. Descriptive statistics and correlations between the task characteristic measures in study 2.

Task Characteristics Measures	*M*	*SD*	1	2	3
1. Strength of cue-answer pairs	0.06	0.08			
2. Number of cue-answer pairs	0.95	0.77	.52**		
3. Number of strong but irrelevant associates	38.25	12.07	−.63*	−.45**	
4. Solution Rate	0.44	0.17	.39**	.42**	-.36**

** $p < 0.01$.
* $p < 0.05$.

Table 4. Final mixed-effects model of the RAT problem-solving performance.

Fixed effects	β (95% CI)	*SE*	*p*	odds ratio (95% CI)
(Intercept)			.001	
Semantic Distance Between the Cues and the Answer				
Number of cue-answer pairs	.19 (.09, .28)	.05	< .001	1.21 (1.09, 1.32)
Strength of cue-answer pairs	.19 (.08, .30)	.06	< .001	1.21 (1.08, 1.35)
Number of Typical but Irrelevant Associates				
Number of strong but irrelevant associates	−.09 (-.19, .01)	.05	.09	0.91 (0.83, 1.01)

Note. 3120 observations, 156 participants, 20 RAT problems per participant. Initial model (random effect only): R^2 = .18, AIC = 3991.7. Final mixed-effects model: R^2 = .21, AIC = 3903.1.

consistent between our data and the normative data. See Table S1 for the solution rate for each RAT problem in our sample.

Generalised linear mixed-effects models with logit link function were constructed to examine the influence of the three task characteristic measures on RAT problem-solving performance (Baayen et al., 2008). The dependent variable was whether or not the RAT problem was solved. Participants were first included in the model as random effects and the three task characteristic measures were entered as fixed effects (see Table 3 for the correlations between these measures). For a mixed-effects model with this specific structure, our sample size (a total of 3120 observations: 156 participants, each solving 20 RAT problems) gives a power of over 80% of detecting significant effects of these task characteristics, if their effects correspond to an effect size *f* of at least 0.175. This power analysis was conducted using the "simr" R package (Green & MacLeod, 2016).

The task characteristic measures were first standardised to allow for comparison of their relative importance. The estimated logit coefficients were transformed to odds ratios for easier interpretation. Table 4 presents the summary of this final model. The inclusion of the task characteristics significantly improved model fit, $\chi(3)^2 = 93.59$, $p < .001$. According to the final model, RAT problem-solving performance was positively associated with the number of cue-answer pairs (odds ratio: 1.21 $p < .001$) and with the strength of cue-answer pairs (odds ratio: 1.21, $p < .001$). As in Study 1, the final model also reported a negative albeit marginally significant relationship between the number of strong but irrelevant associates and RAT problem-solving performance (odd ratio: 0.91, $p = .09$). These results are in keeping with those of Study 1.

The positive effect of the number and strength of cue-answer pairs on RAT problem-solving performance is consistent with previous studies showing that it is difficult to expand the search space to retrieve more remote associates (Marko et al., 2019; Mednick, 1962; Sio et al., 2013). Past studies have also suggested that the presence of irrelevant ideas can impede the ability to solve problems (Smith & Blankenship, 1991). Consistent with this literature, both Study 1 and Study 2 point to a negative, although marginally significant, association between the number of strong but irrelevant associates in the initial problem space and RAT problem-solving performance. A potential reason for this marginality is that our analyses examined the effect of this task characteristic on a wide range of RAT problems. The number of strong but irrelevant associates may have had a larger impact on performance had we focussed only on RAT problems where the cues and the answers are closely associated (i.e., strength of cue-answer pairs > 0). For this type of RAT problem, expanding the initial search space is not that critical for retrieving the answer because the answer is closely associated to some of the cues (Ball & Stevens, 2009; Sio et al., 2013). On such a problem, reaching a solution may depend more on whether an individual could ignore the strong but irrelevant associates that lie between the cues and the answer (Smith, 2003). Thus, the number of these strong but irrelevant associates should be a more critical feature in determining how likely an individual will reach the correct answers of this type of RAT problem.

Additional analyses were conducted to explore the possibility of an interaction between the number of strong but irrelevant associates and the remoteness of the answer on RAT problem difficulty. We re-analysed Study 1's data, conducting the same stepwise regression analysis but only considering RAT problems in which the answer is associated with at least one of the three cues (i.e., strength of cue-answer pairs > 0). Those problems were common—81 out of the 115 RAT problems were of this type. This additional analysis showed that the number of strong but competing ideas became a stronger predictor of RAT problem-solving performance, $\beta = -.36$, $p = .003$ (compared to $\beta = -.20$, $p = .06$, in our original analysis). The small number of RAT problems with remote answers (34 out of the 115 RAT problems have strength of cue-answer pairs $= 0$) precluded us from conducting the same regression analysis on this subset of RAT problems. We thus examined the bivariate relationship between the number of strong and irrelevant associates and RAT problem-solving performance, and found a non-significant relationship between them, $r = .19$, $p = .29$. Together, this pattern of results is in line with the prediction that the number of strong but irrelevant associates would become a stronger predictor of problem difficulty for RAT problems with closer answers.

Generalised linear mixed-effects models with Study 2's data were also constructed to examine if the number of strong but irrelevant associates would interact with the remoteness of the answer (remote answer: strength of cue-answer pairs $= 0$ vs. close answer: strength of cue-answer pairs > 0) on RAT problem-solving performance. There was a significant interaction effect between these two variables ($\beta = -.25$, $p = .02$), suggesting that the negative effect of the number of strong but irrelevant associates was stronger for RAT problems with close answers than for those with remote answers.

In sum, the results of the additional analyses are consistent with our predicted interaction. Given that these analyses were exploratory in nature—which would diminish the interpretability of the reported *p*-values (e.g., Nosek & Lakens, 2014)—we conducted a Bayes factor analysis to quantify the evidence that our data provide in support of our prediction. The Bayes factor analysis revealed moderate to strong support for our predicted interaction (Study 1: $BF_{10} = 15.53$; Study 2: $BF_{10} = 9.37$; see Supplementary Materials for details). To test this predicted interaction more rigorously, we conducted a pre-registered study in which we predicted that the number of strong but irrelevant associates and the strength of the cue-answer pairs would interact with each other for predicting RAT problem-solving performance.

Study 3

In this study, we investigated the interaction between the number of strong but irrelevant associates and the strength of the cue-answer pairs on RAT problem-solving performance. More specifically, we tested whether the negative effect of the number of strong but irrelevant associates on RAT problem-solving performance would become more pronounced when solving RAT problems with stronger cue-answer pairs. This study was preregistered (see protocol at https://osf.io/mqv2p).

Method

Participants

One hundred native English speakers (41 females, 59 males) residing in Canada or the United States were recruited via MTurk. Only MTurk workers with an approval rate of at least 99% were recruited. The mean age was 36.74 years ($SD = 9.84$). 90% of the participants self-reported that they had

Table 5. Descriptive statistics and correlations between the task characteristic measures in Study 3.

Task Characteristic Measures	*M*	*SD*	1	2
1. Strength of cue-answer pairs	0.06	0.08		
2. Number of strong but irrelevant associates	36.73	9.31	-.47*	
3. Solution Rate	.32	.07	.46*	-.49*

*$p < .05$.

a college degree or above. They were paid USD $2 for completing the study, which on average took 20 minutes.

Task

A total of 26 RAT problems were selected from the normative set used in Study 1 (Bowden & Jung-Beeman, 2003b). Supplementary Table S3 lists the selected RAT problems. These 26 RAT problems varied in terms of the strength of cue-answer pairs (ranging from 0 to 0.25) and the number of strong but irrelevant associates (ranging from 16 to 54). Table 5 presents the descriptive statistics and the correlation between these two characteristic measures. To ensure that the RAT problems with strength of cue-answer pairs > 0 were similar in other key task characteristics, all these RAT problems had the same number of cue-answer pairs (value $= 2$).

Procedures

Participants were asked to solve the RAT problems, which were presented once for 30 s in a randomised order. The presentation procedure was identical to Study 2. The study was delivered via Gorilla.sc (Anwyl-Irvine et al., 2020).

Results and discussion

The average solution rate for the RAT problem presented was .32 ($SD =$ .07). The solution rates were positively correlated with those reported in the normative data (Bowden & Jung-Beeman, 2003b) at a moderate level, $r = 0.57$, $p = .003$. See Table S3 for the solution rate for each RAT problem in this study.

Generalised linear mixed-effects models with logit link function were constructed to examine the interaction between the number of strong but irrelevant associates and the strength of cue-answer pairs on RAT problem-solving performance (Baayen et al., 2008). The dependent variable was whether or not the RAT problem was solved. First, a basic model with participants as a random effect was constructed. The two task characteristic measures—the strength of cue-answer pairs and the number of strong but irrelevant associates—were then entered as fixed effects. Lastly, their interaction term was added to the model. As in Study 2, the task characteristic

Table 6. Results of mixed-effect models in Study 3 (without vs. with interaction term).

Variable	β (95% CI)	*SE*	*p*	odds ratio (95% CI)
Model without the interaction term				
(Intercept)			< .001	
Strength of cue-answer pairs	.18 (.06, .31)	.06	< .01	1.20 (1.06, 1.36)
Number of strong but irrelevant associates	−.22 (−.35, −.10)	.06	< .001	0.80 (0.70, 0.90)
Model with the interaction term (Final model)				
(Intercept)			< .001	
Strength of cue-answer pairs	.07 (−.08, .21)	.07	.36	1.07 (0.92, 1.23)
Number of strong but irrelevant associates	−.22 (−.35, −.10)	.06	< .001	0.80 (0.70, 0.90)
Strength of cue-answer pairs*Number of strong but irrelevant associates	−.20 (−.32, −.08)	.06	< .001	0.82 (0.73, 0.92)

Note. 2600 observations, 100 participants, 26 RAT problems per participant. The inclusion of the interaction term significantly improved model fit, $\chi(1)^2 = 10.73$, $p = .002$.
Model without the interaction term: $R^2 = .61$, AIC = 2121.4. Model with the interaction term: $R^2 = .62$, AIC = 2112.7.

measures were standardised to allow for comparison of their relative importance, and the estimated logit coefficients were transformed to odds ratios for easier interpretation.

For a mixed-effects model with this specific structure, our sample size (a total of 2600 observations:100 participants, each solving 26 RAT problems) gives a power of over 80% of detecting a significant interaction, assuming an effect size of $f = 0.175$. This power analysis was conducted using simulation in R ("simr" package, Green & MacLeod, 2016).

The inclusion of the main effects of the associative strength of cue-answer pairs and the number of strong but irrelevant associates improved model fit, $\chi(2)^2 = 36.28$, $p < .001$. Critically, and as predicted in our pre-registration protocol, the inclusion of the interaction term further improved model fit, $\chi(1)^2 = 10.73$, $p = .002$, indicating an interaction between the number of strong but irrelevant associates and the strength of cue-answer pairs. Table 6 presents the summary of the models. Simple slopes analysis was conducted to investigate the interaction (see Table 7 for the results). To examine how the effect of the number of strong but irrelevant associates was moderated by the strength of cue-answer pairs, the regression coefficient of the number of strong but irrelevant associates was estimated when the strength of cue-answer pairs was low (0.77 SD below *M*[1]), medium (*M*) and high (1 *SD* above *M*). Simple slopes analysis indicated a negative effect of the number of strong but irrelevant associates on RAT problem-solving performance, and the negative effect became more pronounced at higher as opposed to lower level of the strength of cue-answer pairs (low: $\beta = -.07$, $p = .40$; medium: $\beta = -.22$, $p < .001$; high: $\beta = -.42$,

[1]Instead of using the conventional value of low (1 *SD* below *M*), we set the value at 0.77 *SD* below *M* because it was the smallest value of the strength of cue-answer pairs in our dataset. Setting the low value at 0.77 SD below *M* was also theoretically meaningful as this represented RAT problems with strength of cue-answer pairs = 0.

Table 7. Results of simple slopes analysis in Study 3.

	Effect of the Predictive Variable on RAT Problem-Solving Performance				
Moderator	Predictive Variable[a]	β	*SE*	*p*	odds ratio
Strength of cue-answer pairs					
Low: −0.77 *SD* (value = 0.00)	Number of strong but irrelevant associates	−.07	.08	.40	0.93
Medium: *M* (value = 0.06)	Number of strong but irrelevant associates	−.22	.06	< .001	0.80
High: +1 *SD* (value = 0.14)	Number of strong but irrelevant associates	−.42	.09	< .001	0.66
Number of strong but irrelevant associates					
Low: −1 *SD* (value = 27.42)	Strength of cue-answer pairs	.27	.07	< .001	1.31
Medium: *M* (value = 36.73)	Strength of cue-answer pairs	.07	.07	.36	1.07
High: +1 *SD* (value = 46.04)	Strength of cue-answer pairs	−.14	.11	.24	0.87

Note. 2600 observations, 100 participants, 26 RAT problems per participant.

.

$p < .001$). This is consistent with our prediction that the number of strong but irrelevant concepts is a critical characteristic in predicting problem-solving performance when solving RAT problems with closer answers. The interaction also implies that RAT problems with close answers are not necessarily easy; they could still be difficult if there is a large number of strong but irrelevant associates (see Figure 1).

Another equally valid way of interpreting the interaction is to consider the effect of the associative strength of cue-answer pairs on RAT problem solving when the number of strong but irrelevant associates was low (1 *SD* below *M*), medium (*M*), and high (1 *SD* above *M*). The associative strength of cue-answer pairs was a significant and positive predictor for RAT performance but only when the number of strong but irrelevant associates was low; its influence on RAT problem-solving performance became statistically non-significant as the number of strong but irrelevant associates increased (low: $\beta = .27$, $p < .001$; medium: $\beta = .07$, $p = .36$; high: $\beta = -.14$, $p = .24$, see Figure 2). This result suggests that the associative strength between the cues and the answer is not always a strong predictor of RAT problem difficulty; rather, its impact depends on the number of strong but irrelevant associates.

In sum, the results confirm the prediction that there is an interaction between the number of strong but irrelevant associates and the strength of cue-answer pairs in predicting RAT problem-solving performance. The pattern of the interaction clarifies the way these characteristics influence the difficulty of RAT problem solving.

General discussion

Three studies were conducted to examine the impact of different task characteristics on RAT problem-solving performance. Study 1 analysed a

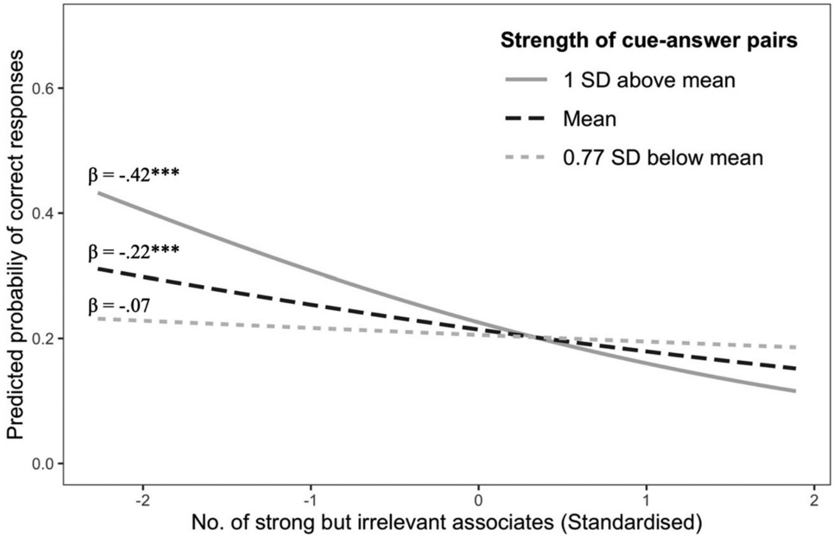

Figure 1. The relationship between the number of strong but irrelevant associates and RAT problem-solving performance as moderated by the strength of cue-answer pairs.
Note. $^{***}p < .001$.

normative dataset and showed that (a) semantic distance between the cues and the answer (measured in terms of the number and strength of cue-answer pairs) and (b) the number of strong but irrelevant associates were key predictors of RAT problem difficulty. Study 2 replicated these findings in a larger and more diverse sample of participants and also suggested that the effect of these factors may interact with one another. Study 3 confirmed this prediction, clarifying the importance and the interdependence of these two factors in determining RAT problem-solving performance.

The positive effect of the number and strength of cue-answer pairs on RAT problem-solving performance is consistent with the current theoretical understanding of creative problem solving (Beaty & Silvia, 2012; Gilhooly et al., 2007; Mednick, 1962). According to this model, individuals first search within the space of concepts related to the problem for an answer, and if no promising answer is found, they expand the search space to retrieve more remote concepts. Similarly, for RAT problem solving, when the answer and the cues are closely related, performing a search of the close associates is likely to be sufficient to find the answer. When the cues and the answer are only remotely associated, individuals must expand the search space to retrieve the remote answer, reducing their chance of discovering the solution within a given time limit.

Our studies reported a significant negative effect of the number of strong but irrelevant associates on RAT problem-solving performance, but only for RAT problems with close answers (i.e., strength of cue-answer pairs

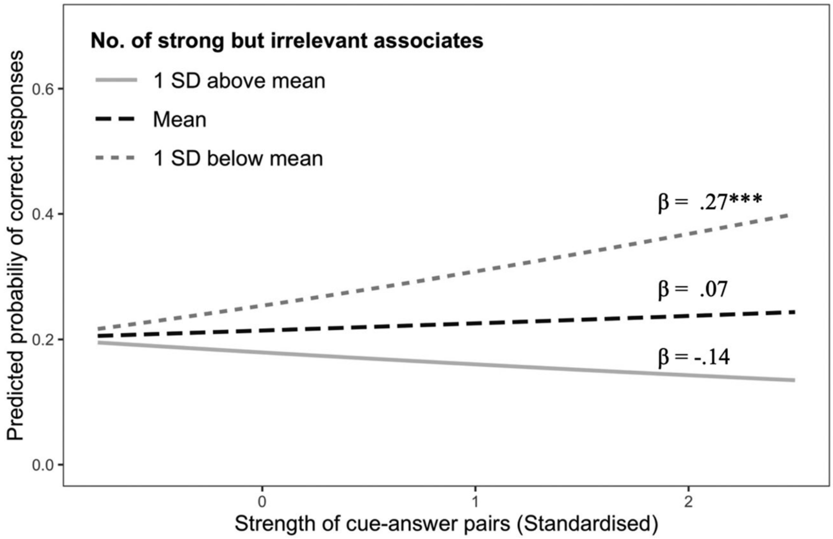

Figure 2. The relationship between the strength of cue-answer pairs and RAT problem-solving performance as moderated by the number of strong but irrelevant associates.
Note. $^{***}p < .001$.

> 0). When solving RAT problems with close answers, the chance of retrieving the correct answer decreased as the amount of strong but irrelevant associates that lie between the problem and the answer increased. This is in line with past studies, showing that the presence of irrelevant associates can prevent the retrieval of the solution and that being able to ignore these associates are critical for successful solution discovery (Luchins & Luchins, 1959; Smith & Blankenship, 1991; Storm et al., 2011). Our studies identify when this factor becomes a critical determinant of RAT problem difficulty. The results also imply that RAT problems with close answers are not always easy to solve as many previous studies have assumed (Davelaar, 2015). The likelihood of solving this type of RAT problem critically depends on the amount of strong but irrelevant associates that lie between the problem and the answer.

When the cues and the answer were remotely associated (i.e., strength of cue-answer pairs $= 0$), the number of strong but irrelevant associates had no significant impact on performance. For RAT problems with remote answers, expanding the search space is needed for reaching the remote answer which is outside the initial space. This search process is likely to be distinct from the process that is critical for solving RAT problems with close answers; otherwise, performance on both types of RAT problems should have been predicted by the same task characteristic—the number of strong but irrelevant associates.

The fact that the number of strong but irrelevant associates is a critical determinant of RAT problem difficulty, but only for RAT problems with close answers, may also explain some inconsistent findings in previous literature concerning the role of inhibition in RAT problem solving. Some studies have suggested that inhibition of irrelevant concepts plays an important role in RAT problem solving (Smith & Blankenship, 1991; Storm et al., 2011), while other studies found no evidence for such relationship (Marko et al., 2019). It is possible that this discrepancy is due to variation in the type of RAT problem used. Indeed, as our findings showed, such discrepancy would be expected if different studies used a different proportion of RAT problems with close answers.

Our results are consistent with past studies suggesting that there is no single optimal search strategy for all RAT problems (Davelaar, 2015; Sio et al., 2013). Our studies extend this line of research by demonstrating that there are different sources of difficulty in RAT problem-solving, each requiring different search strategies. RAT problems could be difficult because of the large semantic distance between the cues and the answer. Conducting a broad search is needed for reaching the remote answer. RAT problems with close answers could still be difficult if there is a large number of irrelevant associates lying between the cues and the answer. Ignoring these competing associates is likely to be a key factor for solution discovery.

The findings that there are multiple determinants of RAT problem difficulty may allow us to better understand the mechanisms of different interventions related to creative problem-solving. RAT problems have been commonly used to measure the effectiveness of different interventions on creative problem-solving, e.g., incubation, task-alternation, and mind-wandering (Lu et al., 2017; Sio & Ormerod, 2009; Zedelius & Schooler, 2015). Although positive effects of these interventions have been reported and different mechanisms have been proposed, strong evidence to discriminate between these proposed mechanisms is still missing. Future studies may compare the effects of these interventions on RAT problems that require different search processes for effective problem-solving, e.g., RAT problems with remote answers vs. RAT problems with comparatively close answers but having a large number of irrelevant associates. The findings could allow us to determine with greater precision the underlying mechanisms of these interventions. These findings could also help identify ways to facilitate performance on other problems that call for similar semantic processes as in RAT problem solving. For example, both RAT problems with remote answers and analogical transfer require the realisation of associations between remote concepts (Cushen & Wiley, 2018; Fu et al., 2013; Gick & Holyoak, 1980). RAT problems with close answers but having a large number of irrelevant associates are similar to other problem-solving domains where disregarding irrelevant information is critical, e.g., overcoming mental set and design fixation (Jansson & Smith, 1991; Viswanathan & Linsey, 2011; Wiley, 1998). Identifying effective interventions for these types of RAT problems would offer important insights into how these creative problems should be approached.

RAT problems are often viewed as semantic retrieval tasks. On the basis of our findings, we suggest that different search processes are required for effective semantic retrieval depending on the remoteness of the target concept as well as the number of competing concepts and that the influence of these two factors are interdependent. The present studies focussed on the relations between task characteristics and search strategies. It is important to note that semantic search processes are also likely to be influenced by non-task characteristics. For instance, previous studies have suggested the importance of attention in semantic processing (Ansburg & Hill, 2003; De Dreu et al., 2012; Zabelina, 2018). Further work could investigate the interaction between task and non-task (e.g., individual differences in attention) characteristics in RAT problem solving. The findings would help develop a comprehensive model that encompasses the influence of both task and non-task factors in semantic search.

Some task characteristics that previous studies have suggested to be predictors of RAT problem difficulty were not significant in our analyses. For example, the number of misleading associates—associates that are closely related to only two out of the three cues—was not a significant predictor of RAT problem-solving performance (the bivariate correlation was also not significant), even though previous studies have suggested that the presence of such misleading associates would induce fixation and thus increase problem difficulty (Sio et al., 2017; Wiley, 1998). One possible explanation for the discrepancy is that the results of these past studies may be due to factors other than the presence of misleading associates. The RAT problems used in the study of Sio et al. (2017) not only contained misleading associates but also linked to a larger number of close but irrelevant associates. Similarly, in Wiley's study (1998), the misleading associates of the RAT problems could misdirect individuals to search for a solution in the area that they were knowledgeable, even though it was not relevant to the solution. It is possible that the difficulty of solving RAT problems with misleading associates in these studies could be attributed to the high number of strong but irrelevant associates, a characteristic that was identified as a relevant predictor of RAT problem difficulty in our studies.

We also did not find a significant impact of the frequency of cue-answer compound words on RAT problem-solving performance. This is in contrast with some previous studies. For example, the study of Olteīeanu and Schultheis (2017) found that RAT problems with low frequency cue-answer compound words were more difficult than those with high frequency cue-answer compound words. It should be noted that only 4% of the RAT problems (5 out of 115 RAT problems) had low frequency cue-answer compound words in our analyses and most of them (73%, 84 out of 115 RAT problems) had high-frequency cue-answer compound words, based on the criteria used in the study of Olteīeanu and Schultheis (2017). The differences in the data range may be one reason that we did not replicate the finding. The non-significant effect of cue-answer compound word frequency in our analyses may also suggest that its impact may diminish after a certain threshold.

Although we expected that RAT problems with a higher level of clustering would be more difficult because switching between clusters is needed to exhaust all the associates in the initial problem space, no evidence was found for this prediction. One possible explanation may be related to the nature of RAT problem solving. The goal of RAT problems is to find one word that is associated with the three cues. The three cues are otherwise unrelated, forming three relatively separate clusters of associates (Davelaar, 2015). Thus, the level of clustering of the initial problem space may be similarly high across RAT problems; the low variability on this measure could be one reason for its non-significant impact on RAT problem-solving performance. It is also important to note that our studies focussed on the impact of the structure of the initial problem space on RAT problem-solving performance. Although searching within the initial problem space is the first and critical stage of problem solving, it is not the only problem-solving stage. In some situations, individuals need to broaden the search to reach relevant concepts that are outside the initial problem space, e.g., solving RAT problems with remote answers. To fully understand RAT problem solving, further studies might explore whether and how the structure of the expanded problem space, i.e., problem space that also includes weak associates (De Deyne et al., 2019), impacts RAT problem difficulty.

Our studies focussed on RAT problem-solving performance. Further studies could measure whether task characteristics also evoke different problem-solving experiences. For example, past studies have suggested that the retrieval of remote associates is done unconsciously (Ansburg & Hill, 2003; Martindale, 1995; Yaniv & Meyer, 1987; Zhong et al., 2008), whereas inhibition of irrelevant associates is often a deliberate process (Altmann & Gray, 2002; Hardt et al., 2013). There has been evidence to suggest that an unconscious problem-solving process could elicit a stronger experience of insight compared to a deliberate problem-solving process (Bowden, 1997; Kounios et al., 2008). In that context, future studies could examine whether RAT problems with remote answers are solved with insight and those with close answers (but with many irrelevant associates) are solved deliberately.

Such a finding would support the idea that these two types of RAT problems require different search strategies. Moreover, the finding would inform us as to why some RAT problems can be solved with insight while some can be solved deliberately, as documented in past studies (Bowden & Jung-Beeman, 2003a; Salvi et al., 2016; Webb et al., 2018).

One may question the generalisability of our findings to other creative thinking tasks due to the convergent nature of RAT problems (i.e., only one correct solution for each item), differing from other creative thinking tasks that involve divergent thinking, e.g., idea generation task (Guilford, 1967; Lee et al., 2014; Lee & Therriault, 2013). Although RAT problems and idea generation tasks may appear different, both can be conceptualised as a search in associative memory requiring individuals to sample ideas from the semantic memory (Beaty & Silvia, 2012; Hass, 2017a, Hass, 2017b). Future studies could examine if the characteristics we identified as relevant predictors of RAT problem difficulty would also be of importance for creative thinking tasks that involve divergent thinking. The findings would help improve the understanding of the general search processes involved in different creative problem-solving situations.

In conclusion, the findings of these three studies suggest that the semantic distance between the problem and the answer, and the number of strong but irrelevant associates in the initial problem space are key factors that determine RAT problem difficulty. The interaction between these factors implies that RAT problems, even with similar solution rates, could be difficult for different reasons and require different search processes. This suggests that there are different types of RAT problems. Future studies using RAT problems as the creative problem-solving tasks should consider selecting RAT problems not only based on their solution rate but also the characteristics of their initial problem space.

Acknowledgments

Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

Funding

This work was partially funded by the Air Force Office of Scientific Research (AFOSR) under grant FA9550-18-1-0088.

References

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

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science*, *13*(1), 27–33. https://doi.org/10.1111/1467-9280.00405

Ansburg, P. I., & Dominowski, R. I. (2000). Promoting insightful problem solving. *The Journal of Creative Behavior*, *34*(1), 30–60. https://doi.org/10.1002/j.2162-6057.2000.tb01201.x

Ansburg, P. I., & Hill, K. (2003). Creative and analytic thinkers differ in their use of attentional resources. *Personality and Individual Differences*, *34*(7), 1141–1152. https://doi.org/10.1016/S0191-8869(02)00104-6

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: an online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Ball, L. J., & Stevens, A. (2009). Evidence for a verbally-based analytic component to insight problem solving. Chapter In N. Taatgen and H. van Rijn (Eds.), *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society*. Austin, Texas: Cognitive Science Society.

Beaty, R., Kenett, Y., Hass, R. W., & Schacter, D. L. (2019). Fanning creative thought: Semantic richness impact divergent thinking. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), Proceedings of the 41st Annual Conference of the Cognitive Science Society (p. 126). Montreal, QB: Cognitive Science Society.

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 309–319. https://doi.org/10.1037/a0029171

Bédard, J., & Chi, M. T. (1975). Expertise. *Current Directions in Psychological Science*, 1(4), 135–139.

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/10.1016/j.intell.2014.05.007

Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(3), 273–281. https://doi.org/10.1037/a0027059

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652–661. https://doi.org/10.1016/j.cognition.2008.05.005

Bowden, E. M. (1997). The effect of reportable and unreportable hints on anagram solution and the Aha! experience. *Consciousness and Cognition*, *6*(4), 545–573. https://doi.org/10.1006/ccog.1997.0325

Bowden, E. M., & Jung-Beeman, M. (2003a). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, *10*(3), 730–737. https://doi.org/10.3758/bf03196539

Bowden, E. M., & Jung-Beeman, M. (2003b). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers : a Journal of the Psychonomic Society, Inc*, *35*(4), 634–639. https://doi.org/10.3758/BF03195543

Brophy, D. R. (1998). Understanding, Measuring, and Enhancing Individual Creative Problem-Solving Efforts. *Creativity Research Journal*, *11*(2), 123–150. https://doi.org/10.1207/s15326934crj1102_4

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*(25), 25. https://doi.org/10.3389/fnhum.2010.00025

Cai, D. J., Mednick, S. A., Harrison, E. M., Kanady, J. C., & Mednick, S. C. (2009). REM, not incubation, improves creativity by priming associative networks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10130–10134. https://doi.org/10.1073/pnas.0900271106

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research, *InterJournal*. *Complex Systems*, *1695*(5), 1–9. http://igraph.org

Cushen, P. J., & Wiley, J. (2018). Both attentional control and the ability to make remote associations aid spontaneous analogical transfer. *Memory & Cognition*, *46*(8), 1398–1412. https://doi.org/10.3758/s13421-018-0845-1

Davelaar, E. J. (2015). Semantic search in the remote associates test. *Topics in Cognitive Science*, *7*(3), 494–512. https://doi.org/10.1111/tops.12146

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. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

De Dreu, C. K. W., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working Memory Benefits Creative Insight, Musical Improvisation, and Original Ideation Through Maintained Task-Focused Attention. *Personality & Social Psychology Bulletin*, *38*(5), 656–669. https://doi.org/10.1177/0146167211435795

Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C., & Wood, K. (2013). The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design*, *135*(2), 021007-1-12. https://doi.org/10.1115/1.4023158

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*(3), 306–355. https://doi.org/10.1016/0010-0285(80)90013-4

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology (London, England: 1953)*, *98*(Pt 4), 611–625. https://doi.org/10.1111/j.2044-8295.2007.tb00467.x

Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, *16*(3), 395–429. https://doi.org/10.1207/s15516709cog1603_3

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind: predicting fluency with PageRank. *Psychological Science*, *18*(12), 1069–1076. https://doi.org/10.1111/j.1467-9280.2007.02027.x

Gruszka, A., & Necka, E. (2002). Priming and Acceptance of Close and Remote Associations by Creative and Less Creative People. *Creativity Research Journal*, *14*(2), 193–205. https://doi.org/10.1207/S15326934CRJ1402_6

Guilford, J. P. (1967). Creativity: Yesterday, today, and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3–14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Gupta, N., Jang, Y., Mednick, S. C., & Huber, D. E. (2012). The Road Not taken: creative solutions require avoidance of high-frequency responses. *Psychological Science*, *23*(3), 288–294. https://doi.org/10.1177/0956797611429710

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends in Cognitive Sciences*, *17*(3), 111–120. https://doi.org/10.1016/j.tics.2013.01.001

Hass, R. W. (2017a). Semantic search during divergent thinking. *Cognition*, *166*, 344–357. https://doi.org/10.1016/j.cognition.2017.05.039

Hass, R. W. (2017b). 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

Hommel, B. (2012). Convergent and divergent operations in cognitive search. In P. M. Todd, T. T. Hills, & T. W. Robbins (Eds.), *Cognitive search: Evolution, algorithms, and the brain* (pp. 221–235). MIT Press.

Humphries, M. D., & Gurney, K. (2008). Network "small-world-ness": a quantitative method for determining canonical network equivalence ": . *PLoS One*, *3*(4), e0002051. https://doi.org/10.1371/journal.pone.0002051

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

Kenett, Y. N. (2018). Going the extra creative mile: The role of semantic distance in creativity – theory, research, and measurement. In *The Cambridge handbook of the neuroscience of creativity* (pp. 233–248). Cambridge University Press. https://doi.org/10.1017/9781316556238.014

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

Kenett, Y. N., Levi, E., Anaki, D., & Faust, M. (2017). The semantic distance task: Quantifying semantic distance with semantic network path length. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(9), 1470–1489. https://doi.org/10.1037/xlm0000391

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 of the United States of America*, *115*(5), 867–872. https://doi.org/10.1073/pnas.1717362115

Kounios, J., Fleck, J. I., Green, D. L., Payne, L., Stevenson, J. L., Bowden, E. M., & Jung-Beeman, M. (2008). The origins of insight in resting-state brain activity. *Neuropsychologia*, *46*(1), 281–291. https://doi.org/10.1016/j.neuropsychologia.2007.07.013

Landau, J. D., & Lehr, D. P. (2004). Conformity to Experimenter-Provided Examples: Will People use an Unusual Feature? *The Journal of Creative Behavior*, *38*(3), 180–191. https://doi.org/10.1002/j.2162-6057.2004.tb01239.x

Latora, V., & Marchiori, M. (2001). Efficient Behavior of Small-World Networks. *Phys Rev Lett*, *87*(19), 198701. https://doi.org/10.1103/physrevlett.87.198701

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*(5), 306–320. https://doi.org/10.1016/j.intell.2013.04.008

Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence? Examining internal and external structure validity evidence of the Remote Associates Test. *Psychology of Aesthetics, Creativity, and the Arts*, 8(4), 446.

Levin, I. (1978). Creativity and two modes of associative fluency: Chains and stars 1. *Journal of Personality*, *46*(3), 426–437. https://doi.org/10.1111/j.1467-6494.1978.tb01010.x

Lu, J. G., Akinola, M., & Mason, M. F. (2017). Switching on" creativity: Task switching can increase creativity by reducing cognitive fixation. *Organizational Behavior and Human Decision Processes*, *139*, 63–75. https://doi.org/10.1016/j.obhdp.2017.01.005

Luchins, A. S., & Luchins, E. H. (1959). *Rigidity of behavior: A variational approach to the effect of Einstellung*. Univer. Oregon Press.

MacGregor, J. N., & Cunningham, J. B. (2008). Rebus puzzles as insight problems. *Behavior Research Methods*, *40*(1), 263–268. https://doi.org/10.3758/brm.40.1.263

Marko, M., Michalko, D., & Riečanský, I. (2019). Remote associates test: An empirical proof of concept. *Behav Res Methods*, *51*(6), 2700–2711. https://doi.org/10.3758/s13428-018-1131-7

Martindale, C. (1995). Creativity and connectionism. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 249–268). The MIT Press.

Mednick, M. T., Mednick, S. A., & Jung, C. C. (1964). Continual association as a function of level of creativity and type of verbal stimulus. *The Journal of Abnormal and Social Psychology*, *69*(5), 511–515. https://doi.org/10.1037/h0041086

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

Mendelsohn, G. A. (1976). Associative and attentional processes in creative performance. *Journal of Personality*, *44*(2), 341–369. https://doi.org/10.1111/j.1467-6494.1976.tb00127.x

Moss, J., Kotovsky, K., & Cagan, J. (2011). The effect of incidental hints when problems are suspended before, during, or after an impasse. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(1), 140–148. https://doi.org/10.1037/a0021206

Nelson, D. L., McEvoy, C. L., Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology*, 45(3), 137–141.

Olteţeanu, A. M., & Falomir, Z. (2015). comRAT-C: A computational compound Remote Associates Test solver based on language data and its comparison to human performance. *Pattern Recognition Letters*, *67*, 81–90. https://doi.org/10.1016/j.patrec.2015.05.015

Olteīeanu, A. M., & Schultheis, H. (2017). What determines creative association? revealing two factors which separately influence the creative process when solving the remote associates test. *The Journal of Creative Behavior*, *53*(3), 389–395. https://doi.org/10.1002/jocb.177

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks*, *31*(2), 155–163. https://doi.org/10.1016/j.socnet.2009.02.002

Paton, B., & Dorst, K. (2011). Briefing and reframing: A situated practice. *Design Studies*, *32*(6), 573–587. https://doi.org/10.1016/j.destud.2011.07.002

Salvi, C., Bricolo, E., Kounios, J., Bowden, E., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & Reasoning*, *22*(4), 443–460. https://doi.org/10.1080/13546783.2016.1141798

Schilling, M. A. (2005). A "Small-World" Network Model of Cognitive Insight. *Creativity Research Journal*, *17*(2), 131–154. https://doi.org/10.1080/10400419.2005.9651475

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120. https://doi.org/10.1037/a0014212

Sio, U. N., Kotovsky, K., & Cagan, J. (2017). Interrupted: The roles of distributed effort and incubation in preventing fixation and generating problem solutions. *Memory & Cognition*, *45*(4), 553–565. https://doi.org/10.3758/s13421-016-0684-x

Sio, U. N., Monaghan, P., & Ormerod, T. (2013). Sleep on it, but only if it is difficult: Effects of sleep on problem solving. *Memory & Cognition*, *41*(2), 159–166. https://doi.org/10.3758/s13421-012-0256-7

Smith, K. A., & Vul, E. (2015). The role of sequential dependence in creative Semantic Search. *Topics in Cognitive Science*, *7*(3), 543–546. https://doi.org/10.1111/tops.12152

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the Remote Associates Test. *Cognition*, *128*(1), 64–75. doi:10.1016/j.cognition.2013.03.00

Smith, S. M. (2003). The constraining effects of initial ideas. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration*. (pp. 15–31). Oxford University Press.

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. https://doi.org/10.2307/1422851

Storm, B. C., Angello, G., & Bjork, E. L. (2011). Thinking can cause forgetting: Memory dynamics in creative problem solving. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(5), 1287–1293. https://doi.org/10.1037/a0023921

Threadgold, E., Marsh, J. E., McLatchie, N., & Ball, L. J. (2019). Background music stints creativity: Evidence from compound remote associate tasks. *Applied Cognitive Psychology*, *33*(5), 873–888. . https://doi.org/10.1002/acp.3532

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology (2006)*, *64*(3), 447–466. https://doi.org/10.1080/17470218.2010.505292

Viswanathan, V., & Linsey, J. (2011). Understanding fixation: A study on the role of expertise. In *DS 68-7: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design*, Vol. 7: Human Behaviour in Design, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011 (pp. 309–319).

Viswanathan, V. K., & Linsey, J. S. (2013). Role of Sunk Cost in Engineering Idea Generation: An Experimental Investigation. *Journal of Mechanical Design*, *135*(12), 121002-1-12. https://doi.org/10.1115/1.4025290

Webb, M. E., Little, D. R., & Cropper, S. J. (2018). Once more with feeling: Normative data for the aha experience in insight and noninsight problems. *Behavior Research Methods*, *50*(5), 2035–2056. https://doi.org/10.3758/s13428-017-0972-9

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

World Economic Forum (2018). *The future of jobs report 2018*. World Economic Forum.

Wu, C. L., Huang, S. Y., Chen, P. Z., & Chen, H. C. (2020). A Systematic Review of Creativity-Related Studies Applying the Remote Associates Test From 2000 to 2019. *Frontiers in Psychology*, *11*, 573432. https://doi.org/10.3389/fpsyg.2020.573432

Wu, C. L., Zhong, S., & Chen, H. C. (2017). Discriminating the difference between remote and close Association with relation to white-matter structural connectivity. *Plos One*, *11*(10), e0165053. doi:10.1371/journal.pone.0165053

Yaniv, I., & Meyer, D. E. (1987). Activation and meta-cognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *13*(2), 187–205.

Zabelina, D. L. (2018). Attention and creativity. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge handbook of the neuroscience of creativity* (p. 161–179). Cambridge University Press. https://doi.org/10.1017/9781316556238.010

Zedelius, C. M., & Schooler, J. W. (2015). Mind wandering "Ahas" versus mindful reasoning: alternative routes to creative solutions. *Frontiers in psychology*, *6*, 834. https://doi.org/10.3389/fpsyg.2015.00834

Zhong, C.-B., Dijksterhuis, A., & Galinsky, A. D. (2008). The merits of unconscious thought in creativity. *Psychological Science*, *19*(9), 912–918. https://doi.org/10.1111/j.1467-9280.2008.02176.x