The Role of Knowledge Structures &

Motivation on Problem Solving

by

Derek M. Ellis

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2021 by the
Graduate Supervisory Committee:

Gene A. Brewer, Chair
Donald Homa
Chris Blais
Stephen Goldinger

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

Individuals encounter problems daily wherein varying numbers of constraints require delimitation of memory to target goal-satisfying information. Multiply-constrained problems, such as compound remote associates, are commonly used to study this type of problem solving. Since their development, multiply-constrained problems have been theoretically and empirically related to creative thinking, analytical problem solving, insight problem solving, intelligence, and a multitude of other cognitive abilities. Critically, in order to correctly solve a multiply-constrained problem the solver must have the solution available in memory and be able to target and access to that information. Experiment 1 determined that the cue – target relationship affects the likelihood that a problem is solved. Moreover, Experiment 2 identified that the association between cues and targets predicted inter- & intra-individual differences in multiply-constrained problem solving. Lastly, Experiment 3 found monetary incentives failed to improve problem solving performance likely due to knowledge serving as a limiting factor on performance. Additionally, problem solvers were shown to be able to reliably assess the likelihood they would solve a problem. Taken together all three studies demonstrated the importance of knowledge & knowledge structures on problem solving performance.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

The Role of Knowledge Structures & Motivation on Problem Solving

Individuals solve a variety of problems on a daily basis. These problems range in difficulty from relatively easy to incredibly difficult. For example, an easy problem could be choosing what to eat for breakfast given what that person has in the fridge, what they have time to prepare, and their current diet. Conversely, a difficult problem could be when a patient goes to the doctor in order to provide the correct diagnosis, they review their presenting symptomatology, the patient's lab work, and their medical history. While these two problems differ on the surface, they are both a type of multiply-constrained problem which has an underlying structure with a goal state and to arrive there the problem solver uses cues or clues (e.g., symptoms or diet) to delimit the search area in memory to arrive at the correct solution. Critically, in order to solve a problem, the solution must be known or available and accessible or retrievable from memory.

Tulving & Pearlstone (1966) examined retrieval for words contained with lists of varying length. Participants were asked to retrieve words from a previously studied list under both cued and noncued recall conditions. Cued recall elicited better retrieval of studied words than noncued indicating that memory traces were present (available) but not accessible as cueing was required to retrieve the information. Relatedly, effective use of cues and strategy alters performance during verbal fluency tasks. Specifically, Unsworth, Brewer, and Spillers (2013) found that when low working memory participants were given less vague cues, such as "sea animals" rather than "animals", during a verbal fluency task their performance improved. When individuals attempt to solve compound remote associate problems, a problem whereby the solver is given three cue words ("water, skate, cream") and tasked with finding a fourth word that forms a compound word or phrase with each of the cue words (solution: ”ice”), retrieval of possible solutions shows clustering (see Troyer, Moscovitch, & Winocur, 1997) and stronger relations to a single cue rather than a combination of cues (Smith, Huber, & Vul, 2013). Given these tendencies it is important to examine how each of the cues for a compound remote associate item influences problem solving ability. Specifically, the strength of the association between the cue words and targets can be quantified using latent semantic analysis (LSA; Landauer, Foltz, & Laham, 1998). Therefore, our first experiment is going to examine the role the cues play in multiply-constrained problem solving.

Given the importance of retrieval it is necessary for follow-up work to examine how the structure of memory influences problem solving ability. Little work to date has examined the role of memory structure in problem solving, but there is an emerging body of work in the related field of creativity. Creativity and problem solving have a shared history due to the shared nature of convergent and divergent thinking processes. A typical convergent thinking task requires the participant to generate a response to a cue/problem that has a single correct answer (e.g., compound remote associate task). Whereas a divergent thinking task requires the participant to generate multiple possible answers given a cue with no singular correct answer (e.g., alternative uses task). Both convergent and divergent processes are engaged even when the task is designed to emphasize one over the other. Specifically, both tasks require solvers to generate a number of possible responses and a single response is submitted from all possible responses, and then repeated in the case of alternative uses task. Recent evidence has demonstrated that individuals with greater interconnectedness between information stored in memory and stronger associations between remotely associated information in memory perform better on divergent thinking tasks (Beaty, Kenett, & Hass, 2019; Benedek, Kenett, Umdasch, Anaki, Faust, & Neubauer, 2017; Kenett, 2018, 2019; Kenett & Faust, 2019). More interconnectedness allows for an individual to get between different nodes of information in fewer steps. Additionally, when two pieces of information have a stronger association, retrieval of one piece is likely to produce retrieval of the second (Kahana, 1996; Kahana, Howard, & Polyn, 2008). Therefore, given the shared nature of creativity and problem solving, previous demonstrations of shared underlying cognitive abilities (Ellis, Robison, & Brewer, 2021; Lee & Therriault, 2013), semantic network analyses should aid researchers in furthering our understanding of problem solving.

Several researchers have demonstrated that intelligence is critical for problem solving (Ellis et al., 2021; Lee et al., 2013; Chuderski & Jastrzebski, 2018). Critically, a problem solver needs to have solutions already stored in memory to solve a problem. This necessary condition thus points to two critical research questions. First, are problem solvers able to determine above chance whether a solution is stored in memory? Bolte & Goschke (2005) conducted a study where participants were shown cues to possible compound remote associate problems which may have been either solvable or unsolvable. They found that individuals were able to determine, above chance, within 2-seconds if the problem was able to be solved or not (see Topolinksi & Strack, 2009 for a set of follow-up studies). Additionally, problem solvers are able to identify how close they are to arriving at a solution during the problem solving process (Metcalfe & Wiebe, 1987). Second, does knowledge of solutions serve as a limiting factor or is it possible to improve performance through external motivation? Specifically, if participants are provided with external motivation to perform better on a problem solving task can they correctly answer more problems or are they limited by the total number of correct responses stored in memory. Incentivization, through financial reward or other means, have been shown to improve performance in a variety of tasks (Brewer, Lau, Wingert, Ball, & Blais, 2017; Cameron & Pierce, 1994; Utman, 1997). Thus, by manipulating external motivation we can fully examine the roles that motivation and intelligence on problem solving and further identify soft and hard caps on problem solving performance.

To summarize, there are two aspects of memory that are related to an individual's ability to solve a problem: availability and accessibility of target information (see Fig. 1). Therefore the primary aims of these experiments are 1) determine whether cue order affects the likelihood of solving a multiply-constrained problem, 2) determine whether the structure of semantic knowledge in memory is predictive of inter- & intra-individual differences in problem solving ability, 3) determine whether knowledge (availability) operates as a limiting factor to successful retrieval of solutions (accessibility), and 4) determine whether the ability to access problem targets is a cognitive penetrability phenomena (i.e., metacognition).

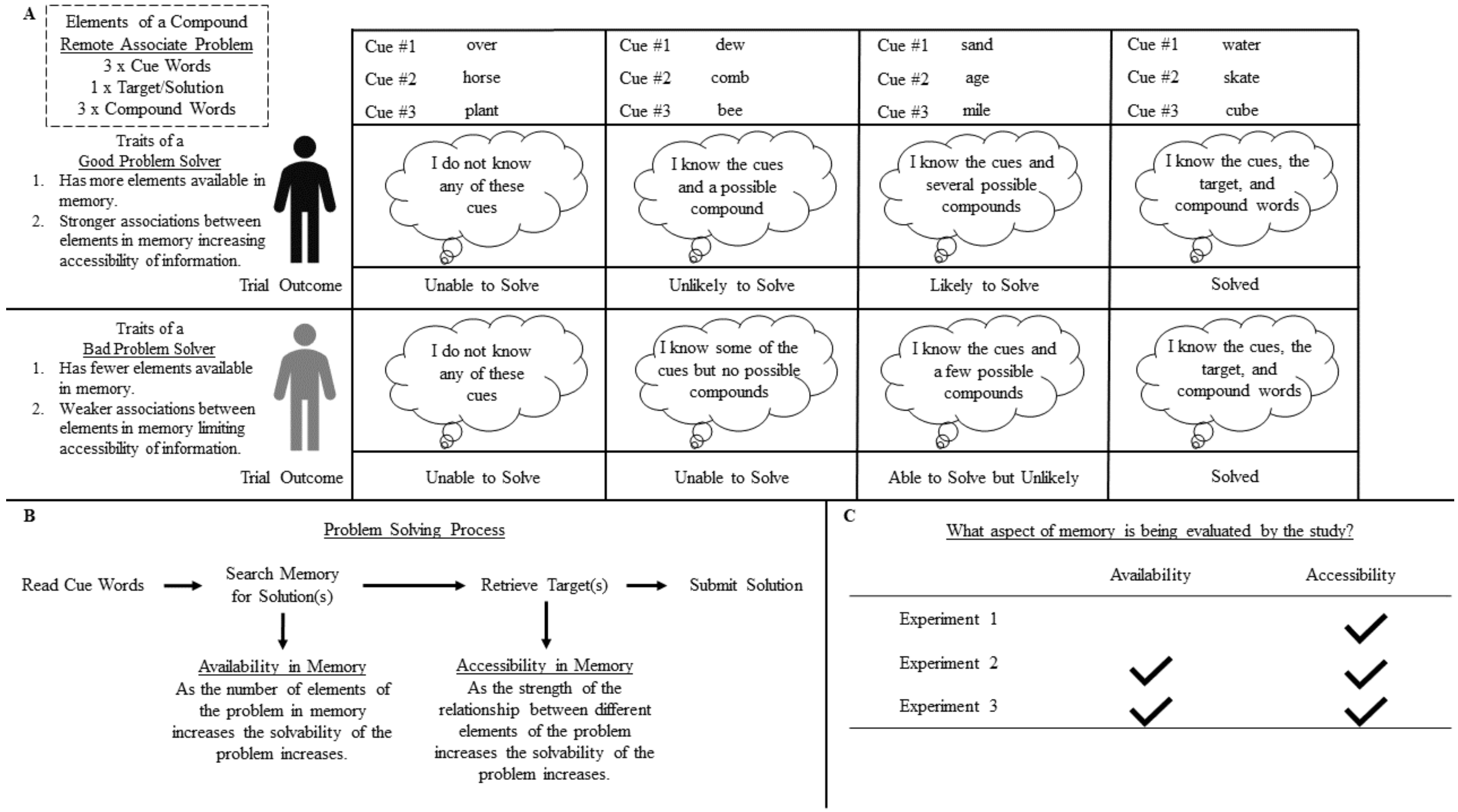

Figure 1. (A) Shows the elements of a compound remote associate problem and how those elements differ between two different problem solvers. Across items the availability of information differs, as well as, the accessibility of information between the good (black) and bad (grey) problem solvers. Overall, the good problem solver has more availability and greater accessibility to elements of the problem. (B) Illustrates the problem solving process and, how and where availability and accessibility play critical roles. (C) Highlights which aspects of memory (availability & accessibility) are evaluated in each of the three experiments.

Chapter 1 (Cue Order)

One of the most commonly used tasks in the problem solving literature is the compound remote associate task (Bowden & Jung-Beeman, 2003). In this task the problem solver is shown three cue words, such as “sand, mile, age”, and asked to find a fourth word that forms a compound word or phrase with each of the cue words (Solution: stone; sandSTONE, mileSTONE, STONEage). To date, it is up to the researcher to determine how the cues are presented on the screen. Therefore, across studies, sometimes the cues are presented left to right with the second cue at central fixation. Alternatively, sometimes the cues are ordered top to bottom with the second cue again at central fixation. Typically, the cues are ordered in the experimental task as they were ordered in the original normative manuscript (Bowden & Jung-Beeman, 2003). Additionally, to date, there is little work establishing which if any cue the solver tends to start with. This lack of specificity creates an interesting dilemma given current findings examining semantic networks and problem solving.

In their seminal work, Collins & Quillian (1969) evaluated the storage of information in semantic memory. They demonstrated that general features of objects may serve as nodes by which another object can be determined to have that feature or be a member of a category (see also Rosche, 1973). Specifically, the retrieval time of semantic information increases based upon the level at which the information is stored. Relatedly, in semantic fluency tasks participants tend to produce a cluster of quick exemplars when they switch to a sub-category (Troyer, Moscovitch, & Winocur, 1997). Follow-up work has demonstrated that when low working memory capacity participants were given specific retrieval cues their performance improved to the same level as high working memory capacity participants (Unsworth, Brewer, & Spillers, 2013). Taken together the collective findings highlight that effective cueing can lead to faster and an overall greater number of retrieved responses. Therefore, for a compound remote associate problem, when the problem solver is shown a cue more related to the target they should be more likely to retrieve it.

Recent evidence has shown that individuals retrieve responses from memory in a sequentially dependent fashion with regards to the strength of the association between the initial cue and each subsequent response (Smith, Huber, & Vul, 2013). Given that the strength between the cues and targets of compound remote associate items can be measured using latent semantic analysis (LSA). Given these two features it is reasonable to believe that whichever cue the solver reads first may shape how they solve that problem. Specifically, if they read the cue that has the strongly association to the target they may be more likely to solve that problem than if they start with the cue that is weakly related. This experiment will examine the effect that cue order has on problem solving ability.

Methods Experiment 1

Participants and Design

An *a priori* power analysis was conducted using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two dependent group means, an effect size (d = .34) derived from a pilot study, and an alpha of .05. A total sample of 56 participants would be required to detect such an effect with 80% power.

We recruited 61 participants from the Arizona State University participant pool and they were given course credit for their participation. Prior to all statistical tests, the data was screened for outliers and 7 participants were removed from all analyses for failing to complete the task as instructed. This study utilized a within-subjects design such that participants attempted an equal number of strongest-to-weakest and weakest-to-strongest compound remote associate problems (alternatively referred to as Cue Order).

Materials

Compound Remote Associate Test. 30 compound remote associate (CRA) items were selected from the Bowden and Jung-Beeman (2003) normed item list (see Appendix). A typical CRA problem requires an individual to search through memory for a target word ("ice") that is semantically related to three cues ("cream, skate, water") and forms a compound word or phrase with each cue. Typically, all three cues are presented on screen at the same time. In our version of the task we presented the cues in sequential order based upon how strong they were semantically related to the target, which was calculated through LSA. Cues could be presented in two orders, strongest association, middle association, and weakest association (strongest-to-weakest) or weakest association, middle association, and strongest association (weakest-to-strongest). Each cue was presented for two seconds before the next cue was presented. After the three cues were presented, they would appear in the order they were presented on the screen simultaneously, from left to right, for the remainder of the trial.

Problems were chosen on the basis they did not have shared cues with other items or a solution that was also a cue for another problem. Moreover, between participants the cues were randomly assigned to one of the two possible cue orders. Participants were given 30 seconds to solve each problem. Two problems were removed from analyses due to transcription errors leading to incorrect presentation of the cues. The dependent variable is the proportion of problems correctly solved.

Results Experiment 1

A paired samples t-test was conducted to compare performance between the strongest-to-weakest and weakest-to-strongest cue order. Performance was significantly different between the two conditions, t (53) = 2.482, p = .016, d = .338. Specifically, our results found performance was better in the strongest-to-weakest condition (M = .264, SD = .181) than weakest-to-strongest (M = .214, SD = .138; see Fig 2.).

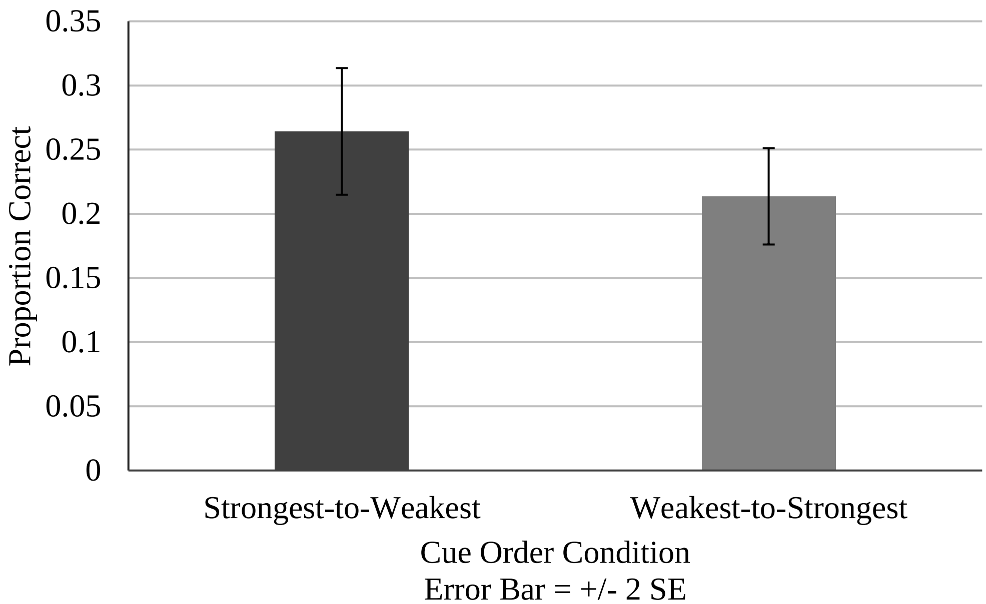

Figure 2. Represents mean performance for the strongest-to-weakest and weakest-to-strongest conditions.

There are several possible control conditions which were not used in this study (e.g., the cue that was not the weakest or the strongest presented first), which can make it hard to interpret the effect of cue order. This analysis will allow us to determine whether sequential presentation of the cues facilitated or impaired problem solving performance relative to normative performance when the three cue words are presented simultaneously. Therefore, in order to determine the direction of the effect, we compared performance in our two conditions against normative data for each item (Ellis & Brewer, In Progress). The overall trend we find is that presenting the cues in sequential order led to overall worse performance relative to known norms. However, for the strongest-to-weakest condition participants showed better performance for number of problems relative to weakest-to-strongest (9 to 5 respectively). Specifically, the strongest-to-weakest condition (*M* = -.031, *SD* = .112) did not perform significantly worse compared to normed performance, *t* (27) = -1.449, *p* = .159, *d* = -.274. However, the weakest-to-strongest condition (*M* = -.085, *SD* = .114) performed significantly worse compared to normed performance, *t* (27) = -3.931, *p* < .001, *d* = -.743 (see Fig. 3).

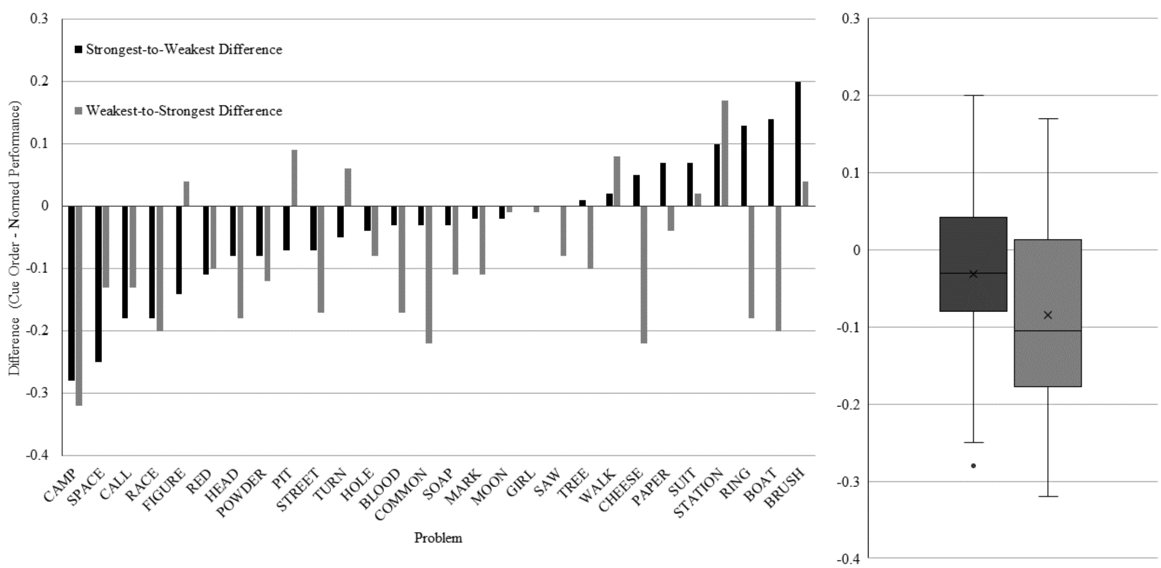

Figure 3. Represents the difference in performance between participant performance and normative data. The x-axis lists the solution word for each problem. The bars are in ascending order based upon the strongest-to-weakest condition. Box-plots show overall performance between strongest-to-weakest and weakest-to-strongest conditions.

Brief Discussion Experiment 1

This experiment examined whether the cue – target relationship affected the likelihood of solving the problem. Specifically, we expected when the cue that was the most semantically related to the target was presented first the problem would be solved more often than when the cue with the weakest relation to target was presented first. The data supported this hypothesis, but a subsequent analysis found that the presentation of the cues one at a time in a sequential order, rather than simultaneously, led to overall worse performance compared to normed performance. However, performance in the strongest-to-weakest condition was not significantly worse than norms but weakest-to-strongest was. Specifically, when participants are shown the cue least semantically related to the target the likelihood of possible targets with stronger associations to the cue is higher and reduces the likelihood that the problem solver will access the correct target. Taken together the findings of this study demonstrate the importance of cue – target relationship in multiply-constrained problem solving. More specifically, the strength of the relationship between the cue and target directly affects the solvers ability to access the target given which cue is selected/read first. However, these findings are based upon an experimental manipulation based upon aggregated population data (i.e., LSA text corpuses). Follow-up is necessary to evaluate cue – target relationships at the individual level which will allow for better determination of who is likely to solve more problems than someone else and which problems an individual is likely to solve.

Chapter 2 (Inter- & Intra- Individual Differences)

To date there is limited research on problem solving and computational models of semantic networks. As highlighted by Mekern, Hommel, & Sjoerds (2019) only five studies examined convergent thinking and three additional studies examined convergent and divergent thinking. Five of those studies examined the compound remote associate task (CRA) by using computational models to better understand the processes and facets of the problem involved (Kajic et al, 2017; Olteteanu & Falomir, 2015; Olteteanu, Falomir, & Freksa, 2018; Olteteanu, Schultheis, & Dyer, 2017; Schatz, Jones, & Laird, 2018). These computational models verified the importance of knowledge and how the strength of the relationship between the cues and targets underlies performance. Specifically, problems with an overall weaker associative or smaller semantic distance between cues and targets are harder to solve than problems with stronger associations between cues and targets. Extending these models into the realm of creative idea generation (divergent thinking) task, such as the alternative uses tasks, researchers began to evaluate the semantic network of individuals. Being able to produce output that matches known human normative problem solving performance is a big step forward for the field, but these efforts emphasize the facets of the problem that change behavior but fail to reveal much about which aspects of an individual's cognitive process and semantic network dictate problem solving.

To better understand what can be learned from mapping semantic networks and related computational models, first we must examine the attributes of the network and what dependent measures can be calculated (see Christensen & Kenett, 2019; Kenett, 2019; & Marupaka, Iyer, & Minai, 2012 for reviews). Much of this discussion will emphasize their importance as it relates to creativity (i.e., generating novel ideas), but we will end with how to transition the primary concepts to problem solving. In studies of semantic networks there are two commonly utilized measures, the clustering coefficient and mean shortest path length. The clustering coefficient represents the level of interconnectedness between nodes in the network. Mean shortest path length is the average number of edges that must be moved along between all pairs of words in the network. Relatedly, the semantic networks of creative individuals and found their semantic networks exhibit many small world tendencies (Beaty, Kenett, & Hass, 2019; Benedek, Kenett, Umdasch, Anaki, Faust, & Neubauer, 2017; Kenett, 2018, 2019; Kenett & Faust, 2019). The general pattern of findings is that individuals who are more creative tend to have higher clustering coefficients (i.e., more interconnectedness) and smaller mean shortest path lengths (i.e., fewer steps between nodes and clusters of nodes). More recently researchers began analyzing creative output using semantic distance measures and it has demonstrated good reliability and predicts human ratings of creative output (Barbot, Hass, & Reiter-Palmon, 2019; Hass, 2017, 2020).

While semantic network mapping has become a prominent feature of studies of intelligence, knowledge structure, and creativity, this procedure does have some noticeable downsides. First, in order to calculate semantic network maps a large volume of data is necessary. This issue is often overcome by aggregating the data from lots of participants and grouping participants based on some secondary factor (e.g., low and high creative ability). Specifically, in these studies participants complete at least one or a few associative fluency tasks. Participants are then grouped on the secondary variable and then you compare the features of the networks between groups. As such, this limitation has made it hard to conduct individual differences research on semantic networks. Zemla et al. (2016) have proposed methods for overcoming this issue but they require participants to complete the associative fluency tasks several times over a longer time interval (i.e., several weeks). Thus, conducting individual differences studies utilizing this alternative method makes it prone to possible missing data or practice effects.

Therefore, our study will utilize a novel measure of semantic/associative distance to overcome these limitations. Specifically, this study utilized compound remote associate problems. Participants were presented with each of the cue words from a set of compound remote associate problems they will attempt to solve. For each individual cue word they will identify a series of word associations. Thus, we can determine for each problem the relative strength between the three cue words and the target. First, we hypothesize that participants who on average have stronger cue - target associations will solve more problems than those with weaker cue - target associations. Second, within an individual, when the cue - target associations are stronger they will be more likely to solve that problem than problems with weaker cue - target associations.

Methods Experiment 2

Participants and Design

An *a priori* power analysis was conducted using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two dependent group means, an effect size (d = 9.55) derived from a pilot study, and an alpha of .05. A total sample of 2 participants will be required to detect such an effect with 80% power.

71 participants were recruited to the study. 2 participants were removed from all subsequent analyses for failing to complete the study as instructed. Therefore, all

analyses were conducted using 69 participants. We utilized a counter-balanced within subjects design to account for possible order effects[1]. Participants completed either the problem solving and word association task and then the remaining task second.

Materials

Compound Remote Association. 20 problems were selected that do not have any overlapping cues or targets. Specifically, this means that a cue or target from one problem did not appear as cues or targets of a different problem. Moreover, targets were not repeated or appear as cues for other problems. Each of the three cues will be used in the word association task. Participants will be with one problem at a time and given 30 seconds to attempt to solve each problem. Problems were randomized between participants. The dependent variable is the proportion of problems correctly solved.

Word Association. The participant was presented with a cue word and asked to generate a single word association, similar to other semantic fluency tasks. Participants were asked to generate a total of 9 associations per word. Participants completed all the associations for a single word before being shown the next cue word. The cues for this task were all of the cue words for the 20 related compound remote associate items. Therefore, participants generated 9 word associations for a total of 60 cue words. For each cue word when participant elicited the target to the related compound remote associate problem their score was the serial position when it was generated (i.e., the associative distance). If the participant failed to generate the target for that specific cue word they were given a score of 10 which is the next best possible position they could have generated the target[1]. The associative distances scores were used to create an average associative distance value for each problem and an overall average associative distance score across all problems.

Results Experiment 2

First, we examined whether associative distance was related to inter-individual differences in problem solving ability. Average associative distance was found to be negatively correlated with the proportion of problems correctly solved, $r = -.528$, $p < .001$ (see Fig. 4). Therefore, participants with overall smaller associative distances (i.e., stronger associations between cues and targets) solved more problems than those with larger associative distances (i.e., weaker associations between cues and targets).

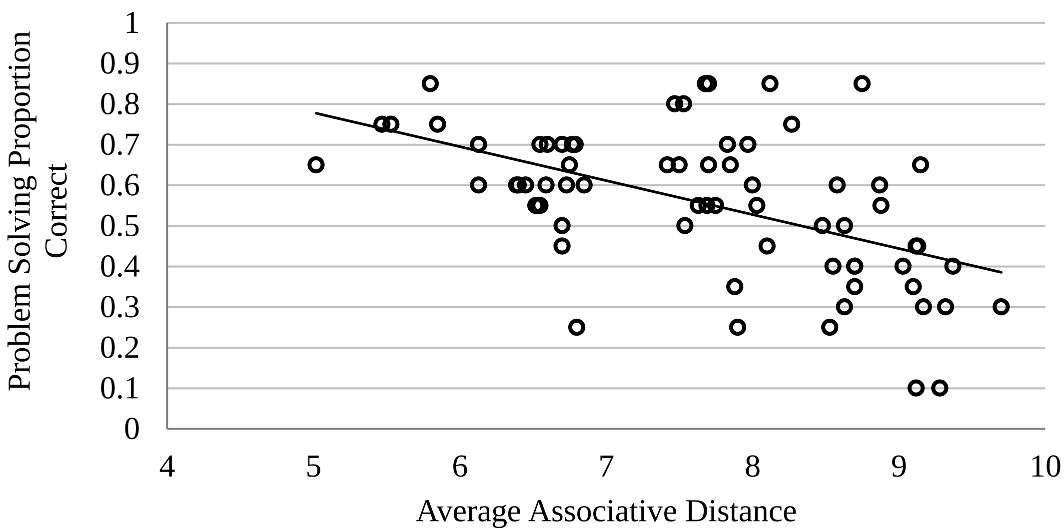

Figure 4. Scatterplot represents the proportion of problems a participant solved in relation to their average associative distance (inter-individual differences). Data points are jittered.

Next, we sought to determine which problems a participant is likely to solve. To evaluate these intra-individual differences a Goodman and Kruskal gamma correlation was calculated comparing the average associative distance with accuracy for each participant independently. Similar to the inter-individual difference analyses we expected to find a negative correlation was expected which would indicate that when the cue – target relationship was stronger (i.e., the target was generated earlier during the word association task) that problem is more likely to be solved than if the cue – target relationship is weaker (i.e., the target was generated later during the word association task). The overall correlation on associative distance (*M* = -.519, *SD* = .294) was found to be significantly different from zero, t (59) = -14.168, $p < .001$, d = 1.83 (see Fig. 5). Giving participants an associative distance score of 10 when they fail to generate the target may be influencing the outcome of this analysis. Therefore, we replicated this analysis using an average associative distance score based upon only cues that the participant elicited the target for. Replicating the previous analysis, the overall correlation on associative distance (*M* = .156, *SD* = .419) was found to be significantly different from zero, t (58) = -2.861, p = .006, d = .373. While replicating the previous analysis only 37 participants demonstrated a negative gamma correlation instead of the 64 participants in the previous analysis.

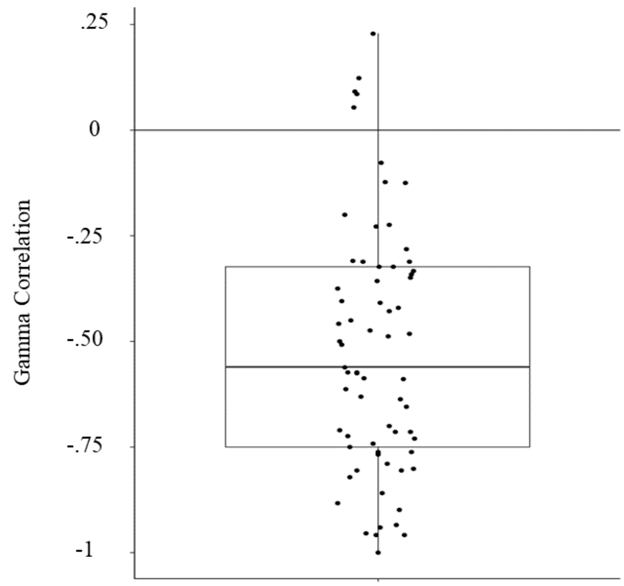

Figure 5. Box-plot showing individual participant Goodman and Kruskal gamma correlations (intra-individual differences). Overall, 64 of 69 participants showed a negative gamma correlations.

Brief Discussion Experiment 2

Previous research has shown that the structure of semantic networks differs between individuals of low and high creative ability (Beaty et al., 2019; Benedek et al., 2017; Kenett, 2018, 2019; Kenett & Faust, 2019). Following up on this related work we quantified the semantic network of participants as they related to a set of multiply-constrained problems. As hypothesized, participants with stronger cue – target associations (i.e., smaller associative distances) solved more problems than those with weaker cue – target associations (i.e., larger associative distances). Similarly, when the cue – target associations is stronger that problem is more likely to be solved than when the cue – target association is weaker. Taken together, the structure of knowledge in semantic memory differs between low and high ability problem solvers. Moreover, at the item by item level we can measure the relationship between cues and targets in semantic memory and predict whether an individual is likely to solve a specific multiply-constrained problem. The findings of experiment 2 provide overwhelming evidence that the accessibility of targets in memory is strongly influenced by the cue – target relationship within an individual. However, given that compound remote associate problems are designed utilizing common English words it is necessary to determine what factors limit an individual's ability to solve problems. For example, an individual may lack the motivation to apply the necessary effort to solve a series of problems. Additionally, given the importance of the cue – target relationship, are problem solvers aware of the likelihood to solve a problem?

Chapter 3 (Metacognition & Incentives)

Problem solving is an important part of everyday life and exist in numerous forms and must be solved in different ways. One example of daily problem solving is choosing an outfit for the day. When choosing an outfit for the day, one must consider multiple limitations or constraints that affect choice. Problems can have a single constraint, such as location, or can have multiple constraints, such as location, season, time of day, activity, and availability of clothing. Solving these multiply-constrained problems employs a two-process approach: the problem solver must first generate a possible solution through semantic search, and then select the solution from the choices generated (Smith, Huber, & Vul 2013). The generation of a solution is a critical step in problem solving and is influenced by pre-solution metacognitive judgements, attention control mediated search processes, and motivation (Nelson & Narens, 1990; Weisberg, 2015; Wieth & Burns, 2006). Therefore, we assessed the effect of motivation on problem solving, including the associated prospective metacognitive monitoring process and the strategy used to find a solution.

The compound remote associate task (CRA) has historically been used to investigate creative thinking (Mednick, 1962) and insight problem solving (Kounios & Beeman, 2009). More recently, the CRA has been used to examine individual differences in multiply-constrained problem solving (Ellis & Brewer, 2018; Ellis, Robison, & Brewer, 2021). The CRA requires a memory-based search to generate a solution and attention control processes to complete the task, both of which are mediated by working memory capacity (Wiley & Jarosz, 2012). Individual differences in working memory capacity was thought explain the majority of individual differences in multiply-constrained problem solving ability (Chuderski & Jastrzebski, 2018). However, recent findings illustrate the importance of crystallized intelligence in multiply-constrained problem solving. Not only do individuals differ in their ability to solve these problems, but also differ in their metacognitive ability to judge the likelihood they will solve these problems.

Problem solving is a multi-step process which begins with an assessment of the problem, the goal, and the likelihood to arrive at the goal. Important in this first step, is whether individuals can accurately assess their own knowledge to determine if a solution can be generated. This judgement is known as a prospective feeling-of-knowing judgement and is important for allocating cognitive resources when searching for a solution (Nelson & Narens, 1990). Previous research has shown that when participants were shown both solvable and unsolvable CRA problems within 2-seconds they could determine above chance what kind of problem it was (solvable vs. unsolvable) (Bolte & Goschke, 2005).The calibration of feeling-of-knowing judgements to an individual's actual ability to solve a problem correctly is a metacognitive-monitoring behavior, wherein an feeling-of-knowing judgment is made, a problem is attempted, and the calibration of the feeling-of-knowing judgment is refined based on the accuracy of original feeling-of-knowing judgement. The relative accuracy of the feeling-of-knowing judgement is an individual's metacognitive ability to predict the likelihood of accuracy in relation to other problems (Dunlosky & Metcalfe, 2009). Relative accuracy is useful in discriminating individual differences in prospective metacognitive judgment accuracy. Individuals vary in the accuracy of their feeling-of-knowing judgments, but overall tend overestimate their ability to solve these problems (Akerman, 2017). Improving individual calibration of this prospective metacognitive judgment can help to improve resource allocation during problem solving tasks. Moreover, poor accessibility and lack of availability of target information in memory should result in a lower reported feeling-of-knowing.

Another way to improve problem solving is by changing the strategy an individual uses to find a solution. Strategy judgements exist on a spectrum between two main strategies: analytic and insight problem solving. Each multiply-constrained problem can be approached from either an analytic approach, an intuitive ‘insight sequence’ approach, or some combination of the two. The analytic thinking approach to problem solving involves continued engagement with the task to actively search for and test each solution. This strategy to problem solving is in direct contrast to the ‘insight sequence’ approach. A solution is reached using an insight approach by actively disengaging from the problem, allowing a non-conscious search of semantic space, and the sudden realization of a solution. While each approach has its advantages for problem solving in general, the insight approach elicits solutions faster, and with greater accuracy when attempting multiply-constrained problems like those in the CRA (Weisberg, 2015). The key difference between these two strategies is the level of focus on the task. While focus is required to complete the task, hyperfocus on the task can actually be detrimental to problem solving ability (Wiley & Jarosz, 2012). The ability to shift strategies and adapt the level of focus needed to complete the task can greatly influence problem solving accuracy. Incentivization may alter which process the solver tends to use. Specifically, the desire to earn more of the incentive (e.g., money) may lead the problem solver to allocate more attentional resources to the task thereby using a more analytical approach.

Motivation is involved at all stages of the problem solving process. Metacognitive judgment may be one mechanism by which motivation is translated into successful problem solving (Nelson & Narens, 1990). Calibration of prospective metacognitive judgments is a self-regulation learning behavior to correct inconsistencies between expected outcomes and actual outcomes. Motivation incentivizes participants to improve self-regulation of learning (Vollmeyer & Rheinberg, 2006). Therefore, if properly motivated an individual will allocate resources to the problem even if a solution is not reached. Moreover, external incentivization to increase the motivation will likely alter solution retrieval processes. Motivation increases goal oriented behaviors like attentional control. Specifically, incentivizing problem solving outcomes will push problem solvers to put more attentional resources into finding the correct answer to a problem. Increased motivation to find a solution enhances attentional control processes and increases the likelihood that problem solvers identify and utilize a successful strategy and ignore an unsuccessful strategy (Sweller, 1983). An increase in attentional focus could encourage greater reliance on an active, analytic search strategy and reduce the reliance on passive insight solution processes. However, a shift to an analytical more attentionally focused search strategy may reduce overall performance for problem sets where insight processes are more likely to be more successful (e.g., compound remote associate problems).

Motivation is one way that problem solving can be influenced, since motivation is an overarching variable throughout the problem solving process. There is a need to understand how motivation impacts all aspects of problem solving. While motivation appears to play an important role in problem solving attempts, understanding how motivation affects both prospective judgements and solution strategies is important to a comprehensive explanation of how motivation assists in problem solving. The goal of this paper is to assess the effects of motivation, metacognitive monitoring, and solution strategies on accuracy in multiply-constrained problem solving. We will accomplish this goal by extrinsically motivating participants with a monetary incentive based on their performance on the CRA. We hypothesize that given the importance of knowledge on problem solving, incentivization will not improve problem solving, but will likely change other aspects of problem solving. For example, given our incentive structure participants in the incentive condition will be less likely to guess and more likely to skip problem they feel they cannot solve. Additionally, we believe participants will be well calibrated to the likelihood they will solve a problem. Lastly, when participants identify a high likelihood of solving a problem they will more likely solve that problem through insight.

Methods Experiment 3

Participants and Design

An *a priori* power analysis was conducted using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two independent group means, a medium effect size (d = .50), and an alpha of .05. A total sample of 126 participants will be required to detect such an effect with 80% power.

One-hundred and thirteen undergraduate students at Arizona State University participated for course credit and were randomly assigned into the incentive or no-incentive condition. No incentive was advertised during participant recruitment. Eight participants were removed from the data analysis due to pre-exposure to the task, three participants were removed from the data analysis for not following instructions, and two participants were removed due to being non-native English speakers. Of the remaining one-hundred native English-speaking undergraduate participants, sixteen were excluded from the incentivized condition for self-reporting no change in motivation as a result of the monetary incentive. The remaining eighty-four participants (N_I=41, N_{NI}=43) were tested in a group laboratory setting for approximately two hours where they completed the compound remote associates task on a computer. Participants rated their prospective feeling-of-knowing judgments before attempting each problem and rated their solution strategy after attempting each CRA problem. The pay structure was designed to encourage maximum problem engagement by rewarding correct solutions, punishing incorrect solutions, and discouraging skipping problems.

Materials

Compound remote associates test. All 144 of the CRA items normed by Bowden & Jung-Beeman (2003) were used. Five new practice problems were used to avoid preexposure to a cue or target word. CRA problems were divided into two groups to limit activation of semantic space: problems with either a shared cue or solution word, and problems with no shared cues or solution words. Participant answers were evaluated to score a response as correct response if it was marked incorrect due to an obvious spelling error (Jakc→ Jack) and non-alphabetic symbols (Jac[k→ Jack). The dependent variable was the proportion of correct solutions entered.

Procedure

Incentive group participants were first informed that they had the chance to earn money based on their performance on the task but could only earn money and would not go into debt for poor performance. The incentive group was then informed of the pay structure: +60 cents for each correct answer, -30 cents for each incorrect answer, and -15 cents for each problem skipped or timed out (i.e., choosing not submit a response or failing to submit a response within the time limit). All participants were informed that the task consisted of four distinct phases per CRA problem: exposure to the cues, FOK rating, time to solve, and strategy rating. Three distinct cue words were presented for 2 seconds (e.g. CRACKER, UNION, RABBIT), after which participants would rate their feeling-of-knowing confidence from 1-6 (i.e. 1 = unlikely to solve, 6 = very likely to solve). Participants were then given 58 seconds to attempt to solve the CRA problem (e.g. JACK) and type their solution or skip the problem. Participants were allowed only one submission per problem. After a solution was submitted, participants were prompted to rate the strategy used to solve the problem from “1” being complete strategy to “4” being complete insight. A strategy rating of “0” could be entered for problems that were skipped or no answer was provided. After a strategy rating was selected, a new set of cues was presented immediately. Participants completed five practice problems and were given feedback at the end of the practice problems before attempting all 144 trials. No feedback was given during the 144 trials. The totals for correct, incorrect, skipped, and timed out problems were presented at the end of all 144 trials, and money earned if applicable.

Results Experiment 3

Given our task there are four possible ways to complete a single problem. Specifically, the participant could provide a correct answer, an incorrect answer, skip the problem (i.e., choose to not submit an answer), and time out or fail to provide an answer within the time limit. Descriptive statistics for each of these outcomes by condition can be found in Table 1. Given these outcomes we conducted separate independent-sample t-test to compare performance between incentive and no incentive conditions. The incentive and no incentive conditions did not differ in the proportion of problem correctly solved, *t* (82) = .070, *p* = .944, *d* = .015, or the proportion of problems timed out, *t* (82) = .737, *p* = .463, *d* = .161. However, the incentive and no incentive conditions did differ in the proportion of problems incorrectly solved, *t* (82) = 2.064, *p* = .042, *d* = .451, and the proportion of problems skipped, *t* (82) = -2.481, *p* = .015, *d* = -.542. Specifically, the no incentive condition (*M* = .413, *SD* = .205) reported more incorrect solutions than the incentive condition (*M* = .327, *SD* = .177). Moreover, the incentive condition (*M* = 270, *SD* = .209**)** skipped more problems than the no incentive condition (*M* = .165, *SD* = .179). Taken together the data indicates that participants were receptive to the incentivization, but it did not improve the number of problems solved.

Table 1.
Descriptive statics for the different problem solving outcomes by condition (Incentive vs. No Incentive)

Condition	Outcome	Min.	Max.	Mean	Std. Dev.	Skew.	Kurt.
No Incentive	Correct	.13	.60	.31	.10	.52	.83
	Incorrect	.09	.83	.41	.20	.16	-1.16
	Skipped	.00	.63	.16	.18	1.15	.61
	Timed-Out	.00	.45	.11	.11	1.347	1.08
Incentive	Correct	.17	.53	.31	.09	.50	-.10
	Incorrect	.03	.72	.33	.18	.16	-.66
	Skipped	.00	.79	.27	.21	.40	-.62
	Timed-Out	.00	.50	.09	.10	1.98	5.82

Following up on the differences found on problem outcomes, incentivization may have altered the process by which solutions are retrieved. For trials when a possible target was submitted (i.e., correct and incorrect trials) we calculated the proportion of trials each solution process (full analytical, partial analytical, partial insight, and full insight) were reported. We conducted a 2 (Group: No Incentive vs. Incentive) x 4 (Solution Process:

full analytical vs. partial analytical vs partial insight vs full insight) mixed ANOVA. There was no main effect of solution process, F (3, 246) = .612, p = .608, partial η^2 = .007, no main effect of group, F (1, 82) = 3.803, p = .055, partial η^2 = .044, and no interaction, F (3, 246) = 1.849, p = .139, partial η^2 = .022 (see Fig. 6). Due to the lack of differences between conditions with regards to overall accuracy the subsequent analyses collapse participants across conditions..

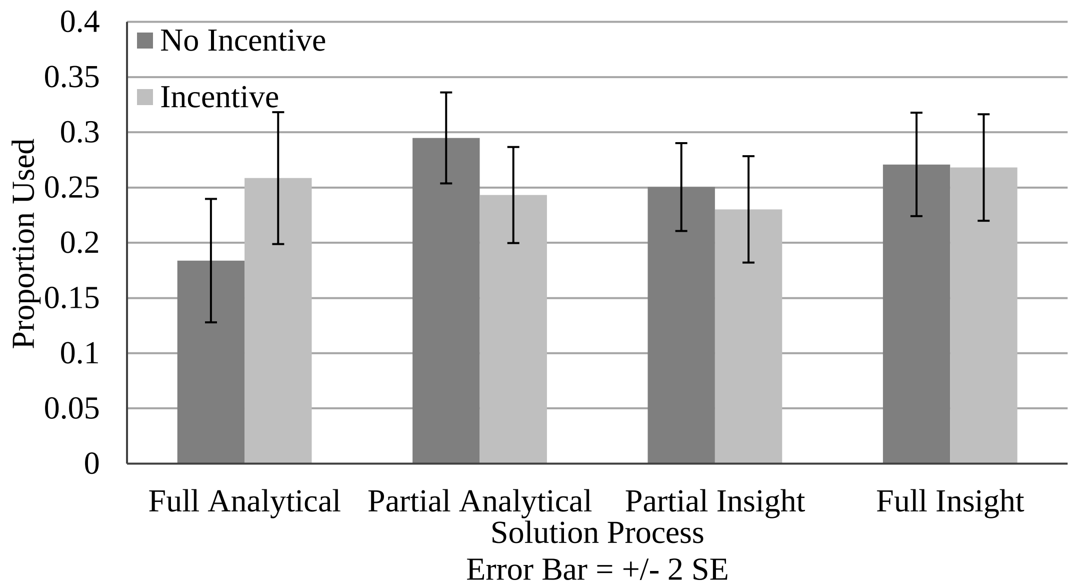

Figure 6. The proportion of trials that each solution process was reported for trials when the participant submitted a possible target.

We then calculated Goodman and Kruskal gamma correlation comparing the feeling-of-knowing rating with accuracy for each participant independently. We expected to find a positive correlation which would indicate that when the participants report a greater feeling-of-knowing they are more likely to solve the problem. This relationship would be indicative of the ease by which the solution is retrieved and the relative certainty of correctness that is associated with insight responses. The overall correlation on feeling-of-knowing (*M* = .402, *SD* = .286) was found to be significantly different from zero, *t* (83) = 12.911, *p* < .001, *d* = 1.409 (see Fig. 7).

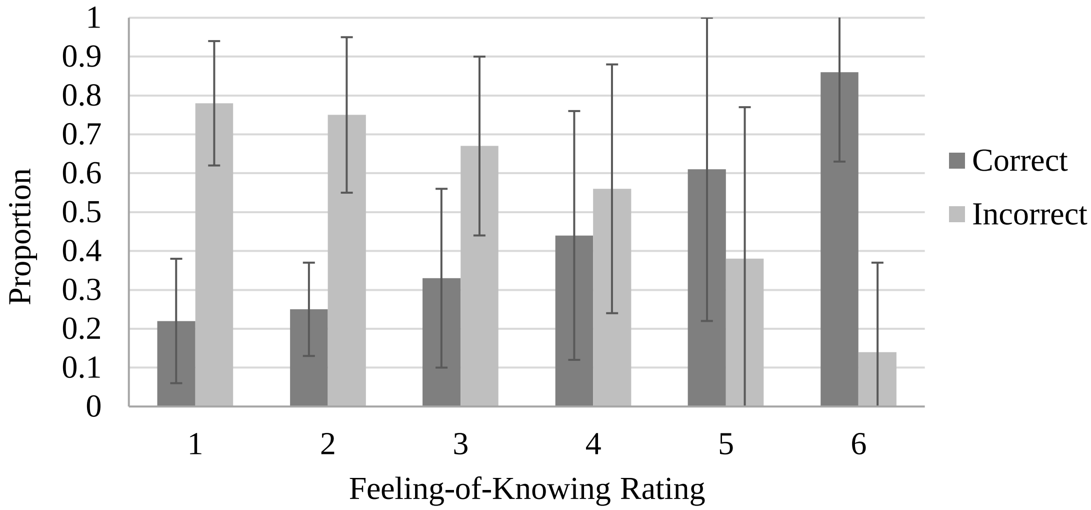

Figure 7. Conditionalized proportion correct/incorrect for each feeling-of-knowing rating. 1: low feeling-of-knowing, 6: high feeling-of-knowing, error bars represent +/- 2 SD.

We then calculated Goodman and Kruskal gamma correlation for correct trials comparing the feeling-of-knowing rating with solution retrieval process (Analytical to Insight) for each participant independently. We expected to find a positive correlation which would indicate that when participants report a greater feeling-of-knowing they more often report having retrieved the answer through insight. The overall correlation on feeling-of-knowing (*M* = .402, *SD* = .286) was found to be significantly different from zero, *t* (83) = 9.329, *p* < .001, *d* = 1.018 (see Fig. 8).

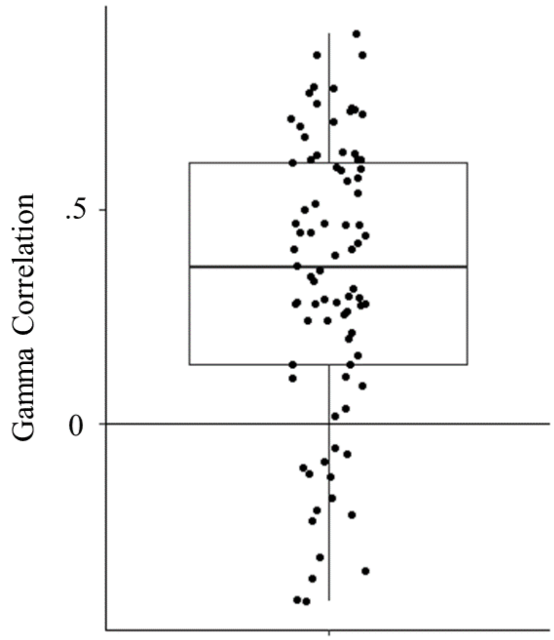

Figure 8. Box-plot showing individual participant Goodman and Kruskal gamma correlations for of feeling-of-knowing by solution process for correct trials. The distribution shows that when participants report a high feeling-of-knowing they often report retrieving the solution through insight.

Brief Discussion Experiment 3

Given the importance of knowledge and structure of knowledge in semantic memory we evaluated the efficacy of monetary incentivization on multiply-constrained problem solving. Additionally, we further examined whether participants were aware of the likelihood to solve a problem (i.e., feeling-of-knowing) and the relation between feelings of knowing and the process by which correct responses are retrieved from memory. We found that participants in the incentive condition were receptive to the monetary incentive, but that it did not result in a greater proportion of correct trials. One possibility for that outcome is that the incentive led to using a more analytical approach which is known to be related to worse performance on multiply-constrained problems

(Salvi, Bricolo, Kounios, Bowden, & Beeman, 2016). However, no statistically meaningful differences were found in solution processes reported between the incentive and no incentive conditions.

As performance between the groups was relatively similar, we collapsed participants across conditions and examined whether they had awareness of the likelihood they would solve a problem. Conceptually replicating previous work, participants are generally aware of which problems are solvable and their likelihood of solving[3]. Specifically, when participants identified a low feeling-of-knowing they solved fewer of those problems relative to when they reported a high feeling-of-knowing. Additionally, we found that when participants correctly solve a problem for which they also had high feeling-of-knowing, they often reported retrieving the correct target through insight.

General Discussion

Experiment 1 demonstrated the importance of the cue – target relationship. When the cue that had the strongest relationship to the target was presented first the problem was solved more often than when the weakest cue was presented first. However, overall performance was worse relative to normed performance. Smith et al. (2013) utilized an external response procedure (i.e., participant reports all retrieved responses, not just the response they believe to be the correct target) and found evidence of sequential dependence of retrieved responses (i.e., subsequent response are related to the previously retrieved response) and that responses tend to be more related to a single cue rather than a combination of cues. However, recent findings have called to question whether their findings are an artifact of the methods used (Howard, Belevski, Eidels, & Dennis, 2020). Specifically, by asking problem solvers to report all retrieved responses the problem solver is being influenced to engage in a serial rather than parallel search. Our findings appear to be more in line with Howard et al. (2020) because if problem solvers are engaging in a serial search during multiply-constrained problem solving then presenting the cues in sequential order should not have impeded performance as much as it did. Given these conflicting findings, the use of eye-tracking, with appropriate cue spacing to reduce the likelihood of all cues being encoded simulataneously or thought probes could be used to investigate which cue a problem solver starts with. Researchers could further ask problem solvers to fixate on the cue they are currently using to solve the problem and can compare semantic related between responses and fixation[4]. A possible follow-up experiment to Experiment 1 and the proposed eye tracking experiment is experimental manipulation of cue direction. Instead of presenting the cues one at a time as we did, all cues are presented simultaneously but the problem solver is instructed to read the cues in a specific order. Comparisons can then be made between directed and undirected reading conditions and may serve as foundational research for a problem solving intervention. Specifically, the difference between a good and bad problem solver could be in large part due to selection of the cue most likely to lead them to solution, which on the aggregate is the cue most semantically related to the target. However, expectations of this intervention should be tempered by the inter- & intra-individual differences found in experiment 2. Specifically, individual differences in the semantic network structure may attenuate any intervention based on semantic measures based upon text corpuses.

In the related field of creativity research, in the last several years the focus has shifted, and begun to emphasize, the differences in the semantic network of low and high creative individuals (Beaty et al., 2019; Benedek et al., 2017; Kenett, 2018, 2019; Kenett & Faust, 2019). While our study lacks the dependent measures used in those studies (e.g., small world tendencies or mean shortest path length), we do conceptually replicate that the relationship between nodes of information in semantic memory underlies differences in abilities. In order to generalize our findings to a similar level of the creativity work, the obvious next step is to evaluate the full semantic network of problem solvers. While the associative distance measure in our study was predictive of inter- & intra-individual differences, we have not demonstrated here and could possibly be unable to generalize or findings to other tasks or possibly other compound remote associate problems. Semantic network measures should provide researchers with the ability to identify the network characteristics of good and bad problem solvers and then generalize those values to a broad set of tasks, some of which may be more ecologically valid when compared to the tasks commonly used in the laboratory (e.g., compound remote associate). We predict, that similar to high creative ability individuals, good problem solvers will exhibit more small world tendencies and smaller mean shortest path length.

In experiment 3 we found overall accuracy between the no incentive and incentive conditions was not statistically different despite participants in the incentive condition eliciting behavior and reporting being receptive to the monetary incentive. One possibility is that our incentive was not significant enough to cause detectable changes in overall performance. Another possibility is that the lack of performance feedback attenuated the overall effect of the incentive or reduced the effect of the incentive over the duration of the study. Alternatively, knowledge and knowledge structures serve as a limiting factor on overall performance. Specifically, a person is unlikely to solve that problem if the solution to a problem is not already stored in memory. Moreover, if the relationship between the cues and targets is missing or weak being able to identify the target, given the cues given during the task, is unlikely as well. Future research should seek to identify which factors (motivation, knowledge/availability of information, and knowledge structure/accessibility) are important and quantify the unique and shared variance amongst them.

Across three studies we investigated the role of availability and accessibility of information in memory affecting the ability to solve multiply-constrained problems. Specific to the compound remote associate (CRA) task, a solver must have the different elements of the problem available in memory to solve the problem. If any element is missing the problem solver cannot with absolute certainty state they will or did solve the problem. Issues related to the number of elements in memory should be classified as failure due to availability. Moreover, just having all the necessary information is insufficient to solve the problem. The elements of the problem must be associated or connected to each other in memory. If all the elements are not associated to one another, the problem solver cannot with absolute certainty state they will or did solve the problem. Issues related to the number of connections between essential elements in memory should be classified as failure due to accessibility. Figure 9 illustrates a few of the possible iterations of how these types of failures might be represented in memory.

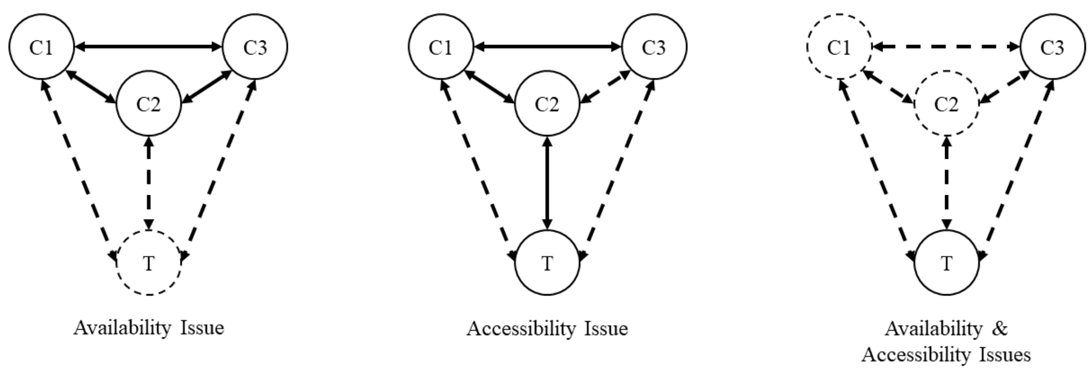

Figure 9. Solid line: present in memory; Dashed line: not present in memory; T: Target; C1, C2, and C3: the three cues for a multiply-constrained problem.

One methods used in problem solving research is the externalized response procedure. In the externalized response procedure, in addition to eliciting a possible solution, the problem solver reports all solutions they generate, in the order they were generated. Pairing an externalized response procedure with semantic network analyses may reveal additional features of good and bad problems solvers. This mixed methodology would allow for better assessment of issues of availability and/or accessibility. Specifically, this will allow researchers to determine how a problem solver “moves” within their semantic space and whether individual differences exist in the type problem solvers make. More precisely, when an individual fails to solve a problem, is that they went down semantic network pathways that were actually less related to the target or did they move down semantic network pathways that were between highly related nodes, but not associated with the target or some combination of the two.

A critical discussion in the problem solving and creativity literature is, what is insight and what cognitive process, if any, does it represent (see Weisberg, 2015 for a review)? One way of evaluating insight is the timeline proceeding insight responses. The core components of this timeline are the problem solver is given a problem state and a goal state they need to reach. Often during the process of trying to achieve the goal state they reach impasse. In order to reach the goal state, they much overcome impasse, often through restructuring, and then reach goal state. Overcoming impasse and becoming aware of the correct response or the steps needed to reach the goal state is associated with an “A-ha” moment. Additionally, the “A-ha” moment is exemplified by the solution coming as if from nowhere. However, previous work has shown that insight is not always the same (Cranford & Moss, 2012). Problem solvers will often report a solution being retrieved through insight without ever having reached impasse. Critically, anytime a solution is brought to conscious awareness it arises out of nonconscious processes. Therefore, there is a distinct possibility that what problem solvers identify as the moment of insight is merely the phenomenological experience of information emerging from nonconscious processes to conscious perception, and if they experience is strong enough or rapid enough it gets identified as insight rather than an analytical process. The gamma correlation analyses in experiment 2 support this possibility. Specifically, higher feelings-of-knowing were associated with greater accuracy and more often paired with solutions having been retrieved through insight. Critically, the feeling-of-knowing rating happens after only 2 seconds of viewing the problem which rules out the problem solving timeline.

In conclusion, knowledge and knowledge structures play important and critical roles in the problem solving process. Measuring an individual's association between cues and targets allows for reliable prediction of inter- & intra-individual differences in multiply-constrained problem solving. Moreover, problem solving behavior can be modified through incentivization, but does not lead to differences in accuracy likely due to knowledge and knowledge structures serving as limiting factors. Specifically, if solutions are not already stored in memory and connections between cues and targets are not present, the likelihood of solving the problem drops significantly. Future research should further examine cue selection, cue preference (i.e., do participants select the cue most semantically related to the target), and other measures of semantic networks which might allow for generalization to other problems and tasks.

FOOTNOTES

[1] Given the counter-balanced design there are two possible task orders 1) word association task then problem solving task or 2) problem solving task then word association task. Participants that completed the problem solving task (*M* = 6.874, *SD* = .912) first had significantly smaller associative distance scores than those who did the word association task first (*M* = 8.513, *SD* = .636), t (67) = 8.583, $p < .001$, $d = 2.068$. However, the groups did not differ on the proportion of problems solved. Thus, we replicated the analyses for each task order independently. Results of these split group analyses are consistent with the main findings reported in Results Experiment 2.

[2] The careful reader may note that by giving the participant a score of 10 when they failed to generate the target word may alter the conclusions of our analyses. Therefore, we rescored the data creating an average associative only using cues for which the target was generated. All analyses were then replicated, and the results were consistent between both sets of analyses.

[3] The previous work mentioned had participants evaluate compound remote associate problems that were either solvable or unsolvable. Participants were able to predict above chance whether a problem was solvable or not (Bolte & Goschke, 2005). Related work has evaluated feelings of knowing more traditionally (Ackerman & Beller, 2017).

[4] However, one risk to this methodology is that by asking participants to fixate on the cue they are currently using will push the solver into using an analytical approach rather than a problem solving process they are more naturally inclined to employ.

REFERENCES

Ackerman, R., & Beller, Y. (2017). Shared and distinct cue utilization for metacognitive judgments during reasoning and memorisation. *Thinking & Reasoning*, *23*(4), 376-408.

Barbot, B., Hass, R. W., & Reiter-Palmon, R. (2019). Creativity assessment in psychological research: (Re) setting the standards. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 233.

Beaty, R. E., Kenett, Y. N., & Hass, R. W. (2019). Fanning creative thought: Semantic richness impacts divergent thinking. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society,* 126-131.

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).

Bolte, A., & Goschke, T. (2005). On the speed of intuition: Intuitive judgements of semantic coherence under different response deadlines. *Memory & Cognition, 33*(7), 1248-1255.

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers, 35*(4), 634-639.

Brewer, G. A., Lau, K. K. H., Wingert, K. M., Ball, B. H., & Blais, C. (2017). Examining depletion theories under conditions of within-task transfer. *Journal of Experimental Psychology: General, 146*(7), 988–1008.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research*, *64*(3), 363–423.

Christensen, A. P., & Kenett, Y. N. (2019). Semantic network analysis (SEMNA): A tutorial on preprocessing, estimating, and analyzing semantic networks. *PsyArXiv*.

Chuderski, A., & Jastrzebski, J. (2018). Much ado about aha!: Insight problem solving is strongly related to working memory capacity and reasoning ability. *Journal of Experimental Psychology: General*, *147*(2), 257-281.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior, 8*(2), 240–247.

Dunlosky, J., & Metcalfe, J. (2009). Chapter 3: Methods and analyses. *Metacognition* (37-59). Los Angeles: SAGE.

Ellis, D. M., Brewer, G. A. (2018). Aiding the Search: Examining Individual Differences in Multiply-Constrained Problem Solving. *Consciousness & Cognition*.

Ellis, D. M., Robison, M. K., & Brewer, G. A. (2021). The cognitive underpinnings of multiply-constrained problem solving. *Journal of Intelligence, 9*(1), 1-25.

Faul, F., Erdfelder, E., Buchner, A., Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods, 41*(4), 1149-60.

Güss, C. D., Burger, M. L., & Dörner, D. (2017). The Role of Motivation in Complex Problem Solving. *Frontiers in Psychology*, *8*, 851.

Kahana, M. J (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24,* 103–109.

Kahana, M.J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. *Psychology, 3*.

Kenett, Y. N. (2018). Investigating creativity from a semantic network perspective. *Exploring Transdisciplinarity in Art and Sciences*, 49-75.

Kenett, Y. N., & Faust, M. (2019). A semantic network cartography of the creative mind. *Trends in Cognitive Sciences*, *23*(4), 271-274.

Kenett, Y. N., Medaglia, J. D., Beaty, R. E., Chen, Q., Betzel, R. F., Thompson-Schill, S. L., & Qiu, J. (2018). Driving the brain towards creativity and intelligence: A network control theory analysis. *Neuropsychologia, 118*(A), 79-90.

Kounios, J., & Beeman, M. (2009). The aha! moment. *Current Directions in Psychological Science*, *18*(4), 210-216.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284.

Lee, C., & Therriault, D. (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,* 306–320.

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), 446-460.

Marupaka, N., Iyer, L. R., & Minair, A. A. (2012). Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. *Neural Networks*, *32*, 147-158.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220-232.

Metcalfe, J. & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*(3), 238-246.

Mekern, V., Hommel, B., & Sjoerds, Z. (2019). Computational models of creativity: A review of single-process and multi-process recent approaches to demystify creative cognition. *Current Opinion in Behavioral Sciences, 27,* 47–54.

Nelson, T.O. & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G.H. Bower (Ed). *The Psychology of Learning and Motivation*, *26*, 125-173. New York: Academic Press

Olteteanu, A., & 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*(1), 81-90.

Olteteanu, A., Falomir, Z., & Freksa, C. (2018). Artificial cognitive systems that can answer human creativity tests: An approach and two case studies. *IEEE Transactions on Cognitive and Developmental Systems, 10*(2), 469-475.

Olteteanu, A., Schultheis, H., & Dyer, J. B. (2018). Computationally constructing a repository of compound remote associate test items in American English with comRAT-G. *Behavioral Research Methods, 50*(5), 1971-1980.

Rosch, E. (1973). Natural categories, *Cognitive Psychology, 4,* 328-350.

Schatz, J., Jones, S. J., & Laird, J. E. (2018). An architecture approach to modeling the remote associates test. *Proceedings of the 16th ICCM.*

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the remote associates test. *Cognition, 128*, 64-75.

Sweller, J. (1983). Control mechanisms in problem solving. *Memory & Cognition, 11*, 32.

Topolinksi, S., & Strack, F. (2009). Scanning the "Fringe" of consciousness: What is felt and what is not felt in intuitions about semantic coherence. *Consciousness and Cognition, 18*, 608-618.

Troyer, A. K., Moscovitch, M., & Winocur, G (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138-46.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior, 5*(4), 381–391.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: the role of controlled search. Memory & Cognition, 41(2), 242-254.

Utman, C. H. (1997). Performance Effects of Motivational State: A Meta-Analysis. *Personality and Social Psychology Review*, *1*(2), 170–182.

Vollmeyer, R., & Rheinberg, F. (1998). Motivationale Einflu¨ sse auf Erwerb und Anwendung von Wissen in einem computersimulierten System [Motivational influences on the acquisition and application of knowledge in a simulated system]. *Zeitschrift fu¨r Pa¨dagogische Psychologie*, *12*, 11 – 23.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review, 18*, 239–253.

Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning, 21*(1), 5-39.

Wieth, M., Burns, D. B.(2006) Incentives improve performance on both incremental and insight problem solving. *Quarterly Journal of Experimental Psychology,* 59(8), 1378 - 1394.

Wiley, J. & Jarosz, A. F. (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science, 21*, 258-262.

Zemla, J. C., Kenett, Y. N., Jun, K.S., & Austerweil, J. L. (2016). *Proceedings of the 38th Annual Meeting of the Cognitive Sicen Society*, 1907-1912.

APPENDIX

Association between each compound remote associate cue and target measured using latent semantic analysis (LSA). Cue 1 is the cue most semantically related to target & Cue 3 is the least semantically related to target for each problem.

Cue 1	**Cue 2**	**Cue 3**	**Target**	**Cue 1 to Target**	**Cue 2 to Target**	**Cue 3 to Target**
way	sleep	board	walk	0.52	0.23	0.12
palm	house	shoe	tree	0.41	0.15	0.09
hammer	hunter	gear	head	0.35	0.27	0.10
cake	cottage	brick	cheese	0.48	0.34	0.09
sense	place	courtesy	common	0.33	0.24	0.16
fox	peep	man	hole	0.31	0.25	0.12
radio	wagon	break	station	0.59	0.29	0.08
carpet	alert	ink	red	0.24	0.10	0.06
keg	puff	room	powder	0.24	0.13	0.09
shine	struck	beam	moon	0.23	0.10	-0.01
paint	hair	sage	brush	0.41	0.31	0.23
summer	ground	boot	camp	0.32	0.24	0.23
mill	dust	tooth	saw	0.23	0.19	0.07
sweeper	main	light	street	0.25	0.23	0.11
drag	horse	human	race	0.53	0.23	0.21
finger	master	toss	ring	0.36	0.20	0.18
check	question	down	mark	0.27	0.25	0.18
law	business	wet	suit	0.19	0.17	0.17
type	ghost	screen	writer	0.15	0.10	0.07
friend	scout	flower	girl	0.38	0.21	0.14
stick	skate	stick	figure	0.15	0.12	0.07
peach	tar	arm	pit	0.11	0.09	0.06
phone	number	cat	call	0.72	0.28	0.09
row	life	show	boat	0.20	0.08	0.07
signal	coat	pike	turn	0.20	0.19	0.07
dish	hand	opera	soap	0.26	0.16	0.06
capsule	ship	cadet	space	0.64	0.14	0.06
clip	fly	wall	paper	0.65	0.06	0.03
pressure	shot	hound	blood	0.25	0.09	0.04
rush	mine	fish	gold	0.57	0.42	0.03