Remote associates test: An empirical proof of concept

Martin Marko¹ · Drahomír Michalko² · Igor Riečanský^{1,3}

Published online: 3 October 2018
© Psychonomic Society, Inc. 2018

Abstract

Associative processes play a major role in research on human thinking, especially creativity. One of the most influential models emphasizing associative processes in creative thinking was introduced by Mednick (*Psychological Review*, 69, 220–232, 1962), who developed the remote associates test (RAT) as a domain-general measure of individual differences in associative hierarchies. Although S. Mednick's theoretical framework has recently regained much attention, the fundamental psychometric assumptions and underlying cognitive processes involved in the RAT remain controversial. We carried out two studies to evaluate these issues. In the first, a confirmatory factor analysis showed that a single latent factor accounted for the ability to solve RAT problems, despite their psycholinguistic heterogeneity. Subsequent regression analyses indicated that cue–solution associative remoteness substantially determined the difficulty of RAT problems, accounting for about 80% of variance. In the second study we used a newly developed associative chain test (ACT), which assesses lexical–semantic and executive measures during associative processing. We found that performance on the RAT was related to lexical–semantic (higher response remoteness and lower response commonness) but not to executive (response inhibition and switching) ACT measures. Overall, our findings indicate that the RAT reflects a coherent ability to access and combine remote elements in lexical–semantic and associative networks without considerably engaging executive attention. Although the validity and utility of the RAT was supported, we propose that the ACT provides a more complex and fine-grained tool for the assessment of associative processing.

Keywords Semantic memory · Executive functions · Creativity · Research methods · Psychometrics

Although it is common in various research fields (e.g., memory, language, and problem solving), assessing word associations has played an especially major role in research on creative cognition (Campbell, 1960; Eysenck, 1995; Mednick, 1962). A particularly influential model relating associative processes and creative thinking was proposed by Mednick (1962). According to this model, creative ideas are achieved by forming mutually remote associative elements into new and useful combinations. These

elements are organized in lexical–semantic and associative structures termed “associative hierarchies.” S. Mednick proposed that people with steeper associative hierarchies are more restricted to a few strongly associated ideas (e.g., common and conventional), whereas people with flatter associative hierarchies are able to access and combine a wider range of weakly related concepts and thus to produce novelty (Mednick, 1962). To operationally define this concept, S. Mednick developed the remote associates test (RAT), which requires the solver to come up with a meaningful link (word association) that mediates three seemingly unrelated cues (e.g., *Same–Tennis–Head* is mediated by *Match*). Given that people with flatter associative hierarchies can generate a larger variety of associative responses to the cue words, including unusual and remote associations, they are more likely to find the missing mediating link and are thus able to solve the test problems more easily than are people with steeper associative hierarchies (Mednick, 1962).

Despite a recent increase in use of the RAT and interest in S. Mednick's conceptual framework (Kenett, Anaki, & Faust, 2014; Kenett et al., 2018; Kounios & Beeman, 2014; Marko & Riečanský, 2018; Marupaka, Iyer, & Minai, 2012; Wu, Zhong, & Chen, 2016), several fundamental assumptions related to both

---

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-018-1131-7>) contains supplementary material, which is available to authorized users.

✉ Igor Riečanský
igor.riecansky@savba.sk

¹ Centre of Experimental Medicine, Department of Behavioural Neuroscience, Slovak Academy of Sciences, Bratislava, Slovakia

² Faculty of Social and Economic Sciences, Institute of Applied Psychology, Comenius University in Bratislava, Bratislava, Slovakia

³ Faculty of Psychology, Department of Basic Psychological Research Methods; Social, Cognitive and Affective Neuroscience Unit, University of Vienna, Vienna, Austria

the RAT and the entire concept of associative hierarchies remain controversial. First, S. Mednick postulated that individual differences in creative ideation can be explained in terms of the steepness or flatness of associative hierarchies (Mednick, 1962), implying that a single latent parameter reflects an individual's "slope" in associative hierarchies. However, the original RAT items are heterogeneous with respect to the way that the cue triplets relate to the solutions (Bowden & Jung-Beeman, 2003; Worthen & Clark, 1971). For example, the cue words *Same*, *Tennis*, *Head* are associated with the response "Match" by synonymy (*same match*), semantic link (*tennis match*), and compounding (*match head*). This heterogeneity applies both within and across items. Moreover, the cue–solution relation can also vary in the degrees of abstractness (e.g., *Apple–Tree* vs. *Humor–Sense*) and figurativeness (e.g., *Star–Planet* vs. *Star–Actress*). Given that processing distinct word relations may employ different cognitive systems (Weiland, Bambini, & Schumacher, 2014; Worthen & Clark, 1971; Wu & Chen, 2017; Xiao, Zhao, Zhang, & Guo, 2012), RAT scores need not necessarily reflect a unitary and coherent cognitive ability.

Second, the theory proposes that individuals with flatter associative hierarchies outperform those with steeper hierarchies in the RAT, because they are able to access and combine remote ideas (Mednick, 1962). This implies that the solutions of RAT problems have to be sufficiently remote from the cues (i.e., a weak, uncommon, or infrequent relation) that individuals with flatter associative hierarchies will have an advantage over individuals who produce close and common associates. Moreover, the concept of associative hierarchies also implies that item remoteness is the principal determinant of item difficulty, so that high test scores reflect better remote associative abilities (Davelaar, 2015). Following this rationale, it is possible that an optimal threshold and variability of item remoteness may be required in order to ensure the validity and sensitivity of the measure. Although the link between item remoteness and other psychometric properties of the RAT (such as item difficulty and item sensitivity) is essential, it lacks empirical confirmation.

Third, empirical tests of S. Mednick's ideas have brought mixed results. A primary role of associative processing was indicated by the finding that individuals with high RAT performance produced greater numbers of associations and maintained relatively high associative speed during a continuous word association task (Mednick, Mednick, & Jung, 1964). These findings are in line with the prediction that individuals with "flat" associative hierarchies can access more (unusual and remote) associations, and therefore preserve higher levels of fluency, even after the dominant responses have been depleted. Also using continuous associations, Benedek and Neubauer (2013) showed that not only fluency, but also the uncommonness of responses, increases with creativity. However, high- and low-creative individuals did not differ in the strengths of their associative responses, indicating that

their lexical–semantic and associative structures (i.e., associative hierarchies) were similar. This conclusion was also reached by Coney and Serna (1995), who failed to find evidence for enhanced priming of remote word pairs in high-creative individuals, as the theory predicts. In contrast to S. Mednick's associative interpretation, some researchers have proposed that higher associative fluency and uncommonness in high-creative individuals may be due to better adaptive control of thought rather than to lexical–semantic processes (Benedek & Neubauer, 2013; Benedek, Panznerer, Jauk, & Neubauer, 2017; White & Shah, 2006). For instance, inhibition and switching have been considered to play important roles in suppressing prepotent responses and overcoming mental set in the RAT (Lee & Theriault, 2013). However, network-analytical techniques have demonstrated that lexical–semantic and associative networks are more interconnected, flexible, and robust for high- than for low-creative individuals (Kenett et al., 2014; Kenett et al., 2018). Such features allow for effective spread of information throughout the semantic network, flexible access to network elements, and the formation of novel conceptual combinations (Marupaka et al., 2012). Taking all these findings together, it remains unclear to what extent performance on the RAT relies on lexical–semantic (associative hierarchies) and executive (cognitive control) systems.

With respect to the abovementioned issues, we approached the theoretical and empirical utility of the RAT in contemporary research in two studies. The first study was a psychometric and methodological evaluation of the heterogeneity and remoteness of RAT problems using both score- and item-level analyses. We tested whether a theoretically heterogeneous set of items could be accounted for by a single latent factor (as is implicitly assumed by the theory; Mednick, 1962) or whether multiple factors are required. Next, we assessed the relationships among psychometric parameters (difficulty and sensitivity) and the various psycholinguistic properties of RAT items. In particular, we evaluated the fundamental assumption that item remoteness (i.e., the associative distance between cues and the respective solution of RAT items) is the main determinant of item difficulty.

In the second study, we evaluated the involvement of lexical–semantic and executive functioning in RAT. For this purpose, we developed a novel associative chain test (ACT). In this test, participants continuously generated word chains following a set of rules. The rules were selected a priori so that performance on lexical–semantic (response commonness and remoteness), executive (response inhibition and switching), and combined lexical–semantic and executive (i.e., response initiation and fluency) measures could be assessed. This fine-grained analysis of associative processing was used to elucidate whether performance of the RAT reflects individual differences in associative hierarchies (i.e., the continuity, uniqueness, and remoteness of associations) or engages executive

processes that underlie the ability to initiate, suppress, and switch among concurrent and possibly interfering associative responses.

Study 1: Psychometric evaluation of RAT problems

Method

Participants A sample of 203 healthy young adults was recruited from among undergraduate university students to complete the RAT. Three of the participants did not finish the testing. The final sample consisted of 200 participants (86 men and 114 women) between 18 to 28 years of age ($M = 20.5$, $SD = 3.8$). Another group of 104 students (38 men and 66 women) 18 to 29 years of age ($M = 21.4$, $SD = 4.3$) was recruited to provide free associative responses that were used to derive the associative remoteness of the RAT items (see below). We found no statistically significant difference in age, $t(302) = 1.87$, $p = .062$, $d = 0.22$, or sex, $\chi^2(1, N = 304) = 1.18$, $p = .277$, $r = .062$, between the two groups.

Remote associates test A Slovak version of the RAT had been developed following previous pilot studies. In these studies, a larger set of RAT items was created and tested among university students in order to obtain responses (multiple solving attempts were allowed), solving times, and the basic psychometric properties (e.g., difficulty and sensitivity) of the items. On the basis of these data, RAT problems with multiple possible solutions, polychoric item-total correlations of $< .30$, and solving probabilities of $< 5\%$ were excluded. The retained items were subsequently mixed with new RAT items and tested again. This process was repeated several times, resulting in a set of 77 items.

On the basis of psychometric analyses of the pilot RAT problems, the 24 items with the highest internal consistency (Cronbach's $\alpha > .80$), appropriate solving probability ($> 20\%$ and $< 90\%$, centered around 50% to 60%), and optimal solving duration (approximately 10 min) were selected as the final RAT item set. This set was administered in the present study (see the [supplementary online materials](#) for more details). The cue words were displayed on a computer screen in one row. The time limit to solve each problem was 25 s. After reaching a solution, participants pressed a keyboard button and orally reported the solution to the experimenter. If the response was correct, the next RAT item was presented. Otherwise, participant continued solving the problem until the time limit had elapsed. This procedure ensured an equal solving time limit for each item and that no item was skipped or overlooked by participants. At the beginning, six practice items (trials) with no time limit were provided. Each item was assessed for

accuracy. The number of correct responses provided within the time limit was used as the performance index in the statistical analyses.

RAT item parameters The link between psycholinguistic and psychometric item parameters was investigated using the set of 24 RAT items and an additional 53 RAT items, which were selected from the pilot studies but not included in the final test due to the constrained test length and item difficulty range. These additional items were included to increase the range of values of the assessed item parameters (see Table 1), provide a more detailed description of their associations (see Fig. 1), and achieve adequate statistical power. The two sets of items were not significantly different in their assessed psycholinguistic and psychometric item parameters (see Table 1), as evaluated by independent-sample t tests, ($t < 1.13$, $p > .26$, $d < 0.26$).

The psycholinguistic properties of the items were evaluated by four independent raters. After a short training, the raters were instructed to assess all cue–solution pairs of the 77 RAT items for *abstractness* (0–4, where 0 indicates a concrete relation—e.g., *Apple–Tree*—and 4 a highly abstract relation—e.g., *Humor–Sense*), *figurativeness* (0–4, where 0 indicates a literal relation—e.g., *Star–Dust*—and 4 a figurative relation—e.g., *Star–Actress*), and *polysemy* (1–3, indicating the number of distinct meanings by which the solution links to the tree cues; e.g., in the problem *Same–Tennis–Head*, the solution *Match* links to three different meanings). Also, the raters assessed each cue–solution relation to indicate whether the relation was *conceptual* (i.e., the cue was a feature or an instance of the solution, or vice versa; e.g., *String–Violin* or *Music–Art*), *functional* (e.g., *Needle–Thread*), *syntagmatic* (i.e., fixed expressions, phrases, or compound words; e.g., *Fall–Line* or *Life–Time*), or the two words were *synonyms* (e.g., *Sense–Meaning*). The *heterogeneity* of an item was then defined as the number of distinct relation types within the item (i.e., ranging from 1 to 3). For each item, the ratings were averaged across raters. Intraclass correlations indicated acceptable interrater reliability for all four psycholinguistic parameters ($ICC > .71$). Furthermore, two item remoteness measures were estimated. The *syntagmatic remoteness* of the cue–solution pairs was derived from the Slovak National Corpus (prim-7.0-public-all, created in 2016, which includes 1.25×10^9 tokens, 5.8×10^6 words, and 3.7×10^6 lemmas) using word frequencies and collocations in a modified *logDice* formula (see the [supplementary online materials](#) for details). The syntagmatic remoteness of each item was calculated by averaging the three respective cue–solution *logDice* values into one score. The *associative remoteness* of cue–solution pairs was derived from free associations provided by an independent group of 104 participants. The participants were asked to produce five associations that first came to mind in response to the cues. On the basis of these responses, the associative distance of cue–solution pairs was estimated as 1 minus the

Table 1 Descriptive statistics of the RAT item parameters

Item parameter	All items ($n = 77$)				Test items ($n = 24$)			
	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Abstractness ^a	0.00	4.00	1.26	1.07	0.00	4.00	1.26	1.10
Figurativeness ^a	0.00	3.25	0.81	0.72	0.00	2.18	0.79	0.58
Heterogeneity ^a	1.00	3.00	1.98	0.60	1.00	3.00	2.06	0.66
Polysemy ^a	1.00	3.00	2.05	0.77	1.00	3.00	2.09	0.74
Syntagmatic remoteness ^b	4.99	12.44	7.89	1.39	5.62	10.24	7.83	1.38
Associative remoteness ^c	.29	.99	.73	.17	.49	.93	.73	.09
Difficulty ^d	.02	.95	.41	.24	.10	.72	.37	.19

^a Rater-based estimate. ^b Corpus-based estimate. ^c Scores are derived from a bank of free associative responses in participants, $N = 104$. ^d Complement to the solving probability ($1 - P_{\text{correct}}$)

relative frequency of the cases in which a cue evoked the appropriate solution associate ($1 - P_{\text{solution|cue}}$, with higher values indicating higher associative remoteness). The associative remoteness of each item was calculated by averaging the three respective cue–solution distance values into one score. Finally, *item difficulty* was estimated as one minus the proportion of examinees who provided the correct answer. Descriptive statistics for each item parameter are provided in Table 1.

Data analysis The unidimensionality of the 24-item RAT (i.e., whether RAT items reflect a single construct) was tested via confirmatory factor analysis (CFA) using weighted least squares means and a variance-adjusted estimator (WLSMV) in the Mplus 6 software. The fit of the one-factor structure was evaluated with the chi-square test (χ^2), comparative fit index (CFI), Tucker–Lewis index (TLI), and the root mean square error of approximation (RMSEA). Two CFA models were estimated—(i) a less restrictive congeneric model (where factor loadings were allowed to vary) and (ii) a *tau*-equivalent model (where all indicators were restricted to have equal factor loadings)—and were subsequently compared using a corrected χ^2 difference test ($\Delta\chi^2$). Furthermore, the internal consistency of the scale was estimated using the ordinal version of Cronbach’s α and McDonald’s ω (Zumbo, Gadermann, & Zeisser, 2007).

The associations between the item difficulty and rater-based psycholinguistic parameters were assessed using Pearson correlations (two-tailed) and forward stepwise linear regression. The hypothesized relationship between item remoteness and item difficulty was assessed via regression analyses.

Results

The CFA analyses showed that both the congeneric and *tau*-equivalent solutions achieved acceptable fits of the indices (Table 2). A chi-square test showed that imposing equal factor

loadings did not lead to any significant decrease in model fit, $\Delta\chi^2(23) = 24.55$, $p = .374$. Factor loadings of the constrained model were estimated to $\lambda = .535$ (all significant at $p < .001$). Consistency indices of the test scores were also acceptable ($\alpha = .90$, $\omega = .90$).

The Pearson correlation analysis showed that the difficulty of the items was positively associated with item figurativeness, $r(75) = .277$, $p = .015$; item heterogeneity, $r(75) = .249$, $p = .029$; and item polysemy, $r(75) = .290$, $p = .011$; but that the correlation between item difficulty and item abstractness was not statistically significant, $r(75) = .202$, $p = .079$. A linear regression using the forward selection method retained only item polysemy as a significant predictor of item difficulty, and adding the other predictors did not significantly increase explained variance.

The association between item difficulty and item remoteness was assessed using linear regression analysis, which showed that both syntagmatic and associative remoteness strongly predicted item difficulty, $F(2, 74) = 140.16$, $p < .001$, $R^2 = .791$ ($b = .053$, $SE = .012$, $p < .001$, and $b = .947$, $SE = .100$, $p < .001$, respectively; note that the predictors were positively correlated: $r = .64$). The relation of syntagmatic and associative remoteness with item difficulty is illustrated in Fig. 1 (panels A and B). A more detailed view using quantile regression analyses showed that both the syntagmatic and associative slope parameters were significant in each quantile (*tau* ranging from .1 to .9 in steps of .05, $p < .001$). However, the relationship between item remoteness and item difficulty was less reliable at lower item difficulty levels (approximately < 20%; see Fig. 1C and D).

Discussion

Our data strongly imply that solving RAT problems recruits a coherent cognitive capacity to retrieve remote associative elements from lexical–semantic networks. The finding that even heterogeneous items can be accounted for by a unitary latent

Fig. 1 Remote associates test (RAT) item difficulty as a function of syntagmatic remoteness (**A**) and associative remoteness (**B**). Black dots represent the test items ($n = 24$), and gray dots represent the additional set of 53 items. Unstandardized (**C**) syntagmatic and (**D**) associative regression slopes for the quantiles of the items' difficulty (from .1 to .9 by .05 step)

Table 2 Goodness-of-fit indices of RAT confirmatory factor analysis models

Model	χ^2	df	χ^2/df	p	CFI	TLI	RMSEA
Congeneric	317.8 ^a	252	1.261	.003	.924	.917	.036 ^b
<i>Tau</i> -equivalent	323.0 ^a	275	1.174	.025	.945	.944	.030 ^b

^a Mean and variance-adjusted chi-square statistics. ^b RMSEA $\leq .05$ probability $> .976$

are also reported. The gray areas depict the lower and upper 95% confidence intervals (CIs) from quantile regression. Dashed horizontal lines depict the lower and upper 95% CIs of the slopes from linear regressions for the remoteness measures

factor contradicts the previous criticism claiming that RAT items with different cue–solution relations (i.e., semantic vs. syntagmatic) employ loosely related processes (Dailey, 1978) and are differently sensitive to remote-associative or creative ability (Worthen & Clark, 1971). Moreover, the test items were equally sensitive to the latent factor (item loadings were equal), providing evidence that none of the psycholinguistic properties and word relation types is psychometrically preferable.

On the other hand, the psycholinguistic properties were associated with item difficulty. Higher item difficulty of polysemous and heterogeneous items could be accounted for by the fact that distinct meanings and relations (i.e., syntagmatic versus semantic) are represented by different lexical–semantic and associative structures (Eddington & Tokowicz, 2015), and thus evoke less-overlapping associative responses. Such items may therefore require more extensive semantic search, especially the search for remote associates. Moreover, given that candidate solutions for RAT problems are generated primarily on the basis of a single cue at a time (Smith, Huber, & Vul, 2013), the response set evoked by one cue may diverge or mislead the associative responses on the subsequent cues, and therefore decrease the likelihood of converging on a suitable solution. Similarly, the comprehension and production of figurative relations involve extracting or forming abstract connections among concepts or features that are typically not related. The processing of highly figurative relations may therefore also put higher demands on the flexible integration of knowledge and semantic representations (Benedek et al., 2014).

Most importantly, our results demonstrate that the difficulty of RAT problems is closely coupled with remoteness—that is, the associative and syntagmatic distance between cues and the respective solutions (these two measures explained almost 80% of item difficulty variance). In line with S. Mednick's model, these findings indicate that the difficulty of RAT items stems from demands on the lexical–semantic search for remote solution candidates. In particular, our data support the pivotal premise that solving difficult problems requires access to uncommon and infrequent lexical–semantic associates, which is a faculty attributed to people with flatter associative hierarchies. On the other hand, retrieving common and frequent responses, which has been attributed to individuals with steep associative hierarchies, may hinder the ability to solve such items (Mednick, 1962). However, the relationship between item associative remoteness and difficulty seems to be less reliable at lower difficulty levels. In this respect, our results suggest that easy items, with a difficulty below .2 (i.e., a mean solving probability higher than 80%) or associative remoteness below .5 (i.e., a mean probability that the item cues will evoke the solution within the first five prepotent responses higher than 50%), may be less appropriate item candidates (see Fig. 1B and D).

Study 2: Lexical–semantic and executive involvement in the RAT

Method

Participants

An a priori statistical power analysis of expected effect size ($r = .40$; Benedek & Neubauer, 2013; Marko & Riečanský,

2018) indicated that at least 37 participants were required in order to reach sufficient statistical power to test the main hypotheses (repeated measures ANOVA and linear multiple regression F tests, $\alpha = .05$ and $1 - \beta = .80$). Following this estimate, we recruited 40 undergraduate university student volunteers to participate in the study. Due to technical problems, two individuals were excluded, and the final sample consisted of 38 students (20 men and 18 women) aged between 19 and 31 years ($M = 23.2$, $SD = 2.8$). Financial compensation was provided for participating in the study. All participants signed informed consent.

Cognitive assessment

Participants were administered three cognitive tasks in random order, using a computer.

Remote Associates Test (RAT) The RAT included 24 problems from the previous study. The time for solving each item was restricted to 25 s. The items were ordered on the basis of difficulty, in ascending order. In line with Study 1, the internal consistency of the test was acceptable ($\alpha = .86$, $\omega = .87$).

Associative chain test (ACT) In this newly developed test, participants continuously generated word chains according to specific rules. In the associate chain, each new response word in the chain was required to be semantically related to the previous one (e.g., *Doctor* [starting word] $\leftarrow^{(A)}$ *Hospital* [1] $\leftarrow^{(A)}$ *Room* [2] $\leftarrow^{(A)}$ *Table* [3] $\leftarrow^{(A)}$ *Pen* [4]). The participants were instructed that entering an unrelated word would be considered as error. In the dissociate chain, each new response word in the chain should *not* be related to the previous one (e.g., *Teacher* [starting word] $\leftarrow^{(D)}$ *Kitchen* [1] $\leftarrow^{(D)}$ *Hockey* [2] $\leftarrow^{(D)}$ *Apple* [3] $\leftarrow^{(D)}$ *Book* [4]). Here, the participants were instructed that producing related words would be scored as errors. In the associate–dissociate chain, participants were asked to deliver associations and dissociations in alternation (e.g., *Phone* [starting word] $\leftarrow^{(A)}$ *Call* [1] $\leftarrow^{(D)}$ *Banana* [2] $\leftarrow^{(A)}$ *Fruit* [3] $\leftarrow^{(D)}$ *Car* [4]). The ACT thus includes two independent variables (factors) with two levels each. The first factor is response relation type, *associative* or *dissociative*, and the second factor is sequence type, *fixed* or *alternating*.

To evaluate performance on the ACT, each response (i.e., each generated word) can be assessed with several measures. The main dependent measures used in our study were *response time* (RT; time required to initiate word responses), *response commonness* (RC; natural logarithm of word frequencies in corpus), and *response remoteness* (RR; based on syntagmatic corpus collocations and the modified *logDice* formula, which was referred to as *syntagmatic remoteness* in Study 1). RC and RR represent lexical–semantic measures.

To assess *inhibition cost*, the RT difference when delivering dissociative (dissociate RT) versus associative (associate RT) responses was calculated (separately for fixed and alternating sequences). Comparably, to assess *switching cost*, the RT difference when delivering responses within alternating versus fixed sequences was calculated (separately for associative and dissociative responses). Inhibition cost and switching cost both indicate executive attentional demands (see the [supplementary online information](#) for more details).

In addition, a separate *remote* associate chain was also introduced to assess the search for remote associates similar to that involved in RAT. Participants were instructed that each new response word in the chain should be remotely associated with the last one. A “remote associate” was described as an uncommon, infrequent, and/or original word associate (e.g., *Dog* [starting word] $\leftarrow^{(RA)}$ *Guard* [1] $\leftarrow^{(RA)}$ *Fort* [2] $\leftarrow^{(RA)}$ *Brick* [3] $\leftarrow^{(RA)}$ *Chalk* [4]). The participants were instructed to deliver a response as remote as possible but to maintain high fluency, and that both strongly related (prepotent) and entirely unrelated words would count as errors.

A short training of each chain was provided before the assessment. Each chain started with the presentation of a randomly selected word from a word list. Thereafter, participant started to continuously produce responses using the computer keyboard until the time had elapsed (we used 60, 90, 150, and 90 s, as the times for associate, dissociate, associate–dissociate, and remote chains, respectively). Participants were instructed to ignore grammatical or typing errors, to maintain fluent word production, and not to repeat the same words within the same chain. In our study, all four chains were generated three times in the same order (yielding a total of 12 associative chains). The duration of the whole test was approximately 20 min.

Hayling sentence completion test (HSCT) This standard neuropsychological test requires participants to complete sentences with either meaningfully fitting words (the so-called *initiation* condition; e.g., “The murderer was sentenced to twenty . . .” completed by “years”) or completely unrelated/incompatible words (*suppression* condition; e.g., “The murderer was sentenced to twenty . . .” completed by “apples”). In both conditions, participants listened to a set of 20 sentences with the final word omitted (the sentence set and condition were counterbalanced). The sentences were recorded in a female voice and presented in random order within each condition. The time and accuracy of each response were assessed (errors were excluded from the analyses). There is convincing evidence that response suppression, in contrast to response initiation, engages cognitive inhibition (Allen et al., 2008; Collette et al., 2001). Therefore, the RT difference between the suppression and initiation conditions was used as a marker

of inhibition processing cost and cognitive control. The three HSCT measures (i.e., response initiation RT, suppression RT, and inhibition cost) were used to evaluate the concurrent and discriminant validity of the corresponding ACT measures (i.e., associate RT, dissociate RT, and inhibition cost, respectively).

Data analysis

The RAT scores of two participants were excluded due to prior knowledge of some RAT solutions. The RT, RC, and RR values of all generated word responses were winsorized (10% quantile two-sided trimming) separately for each individual and condition before computing the average measures used in statistical analyses. Two RT values so averaged (one associate RT in the ACT and one suppression RT in the HSCT) were considered extreme (more than three *SDs* from the mean) and therefore were excluded from the analyses. The rest of the winsorized average measures from the ACT and HSCT were analyzed using repeated measures analysis of variance (ANOVA) models. The association of these measures with RAT performance was evaluated using a linear regression analysis. The *p* values of the ANOVA and regression models were adjusted using sequential Bonferroni correction, as suggested by Cramer et al. (2016). Significance tests of Pearson correlation coefficients used to explore the associations between RAT and ACT measures were not corrected for multiple testing, and therefore should be considered with caution.

Results

HSCT response time

A one-way repeated measures ANOVA comparing RTs between the initiation ($M = 0.464$, $SE = 0.042$) and suppression ($M = 3.177$, $SE = 0.321$) conditions revealed a robust effect of the HSCT condition, $F(1, 36) = 79.16$, $p < .001$, $\eta_p^2 = .687$. The inhibition cost in the HSCT was on average 2.71 s ($SE = 0.31$).

ACT response time

A two-factor repeated measures ANOVA for RTs showed a significant interaction between response type and sequence type, $F(1, 36) = 10.34$, $p = .003$, $\eta_p^2 = .223$. Both main effects were also significant: $F(1, 36) = 113.02$, $p < .001$, $\eta_p^2 = .758$, for response type, and $F(1, 36) = 33.59$, $p < .001$, $\eta_p^2 = .483$, for sequence type (*p* values were adjusted using sequential Bonferroni correction). The dissociative RT was substantially longer than the associative RT, but this difference (i.e., inhibition cost) was bigger in alternating than in fixed sequences (see Fig. 2 for more details). A comparison of RTs between

associative fixed, dissociative fixed, and remote associate chains using a one-way repeated measures ANOVA indicated a significant difference, $F(2, 58.3) = 138.96, p < .001, \eta_p^2 = .794$. Bonferroni-corrected post-hoc t tests revealed that the RT in the remote condition was significantly longer than those in both the associative fixed condition, $t(36) = 14.2, p < .001, d = 2.33$, and the dissociative fixed condition, $t(36) = 9.8, p < .001, d = 1.59$ (see Fig. 2).

ACT response commonness

A two-way repeated measures ANOVA for RC showed a significant main effect of response type, $F(1, 37) = 63.21, p < .001, \eta_p^2 = .631$ (RC was lower in dissociative responses); a nonsignificant main effect of sequence type, $F(1, 37) = 4.71, p = .073, \eta_p^2 = .113$ (it tended to be lower in alternating sequences); and a nonsignificant interaction of the two factors, $F(1, 37) = .28, p = .601, \eta_p^2 = .007$ (p values were adjusted using sequential Bonferroni correction; see Fig. 3). A comparison of RC between the associative fixed, dissociative fixed, and remote conditions using a one-way repeated measures ANOVA indicated a significant difference, $F(2, 74) = 15.62, p < .001, \eta_p^2 = .297$. Bonferroni-corrected post-hoc t tests revealed that the RC of responses was higher in the associative fixed condition than in the dissociative fixed, $t(37) = 5.74, p < .001, d = 0.93$, and remote, $t(37) = 3.90, p < .001, d = 0.63$, conditions, but there was not a statistically significant difference between the dissociative fixed and remote conditions, $t(37) = 1.91, p = .192, d = 0.31$ (see Fig. 3).

ACT response remoteness

A two-way repeated measures ANOVA for RR showed a substantial effect of response type, $F(1, 37) = 2,302.31, p < .001, \eta_p^2 = .984$ (RR was higher in dissociative responses); a

Fig. 2 Mean response times in the associative chain test. Error bars represent 95% CIs

Fig. 3 Mean response commonness in the associative chain test. Error bars represent 95% CIs

significant effect of sequence type, $F(1, 37) = 5.56, p = .024, \eta_p^2 = .131$ (RR was lower in alternating sequences); and a nonsignificant interaction of the two factors, $F(1, 37) = 4.81, p = .035, \eta_p^2 = .115$ (p values were adjusted using sequential Bonferroni correction; see Fig. 4). A comparison of RR between the associative fixed, dissociative fixed, and remote conditions using a one-way repeated measures ANOVA showed a significant difference, $F(2, 74) = 648.15, p < .001, \eta_p^2 = .946$. Bonferroni-corrected post-hoc t tests revealed that RR in the remote condition was significantly higher than in the associative condition, $t(37) = 14.81, p < .001, d = 2.40$, but significantly lower than in the dissociative condition, $t(37) = 21.56, p < .001, d = 3.50$ (see Fig. 4).

Correlation analyses

RAT performance correlated negatively with RT in the associative fixed condition and with response commonness, and positively with response remoteness in the ACT (see Table 3

Fig. 4 Mean response remoteness in the associative chain test. Error bars represent 95% CIs

Table 3 Pearson correlations among the RAT, HSCT, and ACT measures

ACT measures	RAT	HSCT	
		Initiation	Suppression
Response times			
Associative fixed	− .430**	.368*	.226
Associative alternating	− .249	.290	.197
Dissociative fixed	.004	.292	.520**
Dissociative alternating	− .035	.219	.525**
Remote	− .020	.306	.473**
Lexical–semantic			
Commonness ^a	− .387*	.242	− .074
Remoteness ^a	.499**	.074	.462**
Executive (derived)			
Inhibition cost fixed	.237	.223	.659**
Inhibition cost alternating	.156	.074	.607**
Switching cost associative	.085	.203	.320
Switching cost dissociative	− .090	− .118	.117

Bold font indicates statistically significant correlations. * $p < .05$. ** $p < .01$ (not corrected for multiple testing). ^a Computed across all ACT conditions, the intraclass correlation coefficients for response commonness and remoteness across the ACT conditions were $ICC = .71$ and $.73$, respectively. ACT: associative chain test; RAT: remote associates test; HSCT: Hayling sentence completion test

and Fig. 5). The correlation between RAT and remoteness tended to be higher in the dissociative conditions [dissociative fixed: $r(35) = .446$, $p = .006$; dissociative alternating: $r(35) = .629$, $p = .001$] than in the associative conditions [associative fixed: $r(35) = .331$, $p = .049$; associative alternating: $r(35) = .327$, $p = .051$]. However, the pairwise comparisons of correlation coefficients did not show statistically significant differences ($p > .10$). Notably, response commonness and response remoteness were negatively correlated, $r(35) = -.344$, $p = .035$. The executive ACT measures were positively correlated

among each other ($ICC = .890$, averaged inhibition cost measures correlated positively with averaged switching cost measures, $r = .405$, $p = .013$). However, these measures did not show statistically significant associations with the RAT, response commonness, or response remoteness ($p > .115$). Furthermore, neither of the two HSCT measures was significantly correlated with RAT ($p > .182$; see Table 3 for more details).

RAT regression model

Using associative fixed RT, RR, and RC as predictors of RAT performance in a linear regression analysis (backward elimination procedure) showed that RC did not significantly contribute to the model, $t(31) = -1.191$, $p = .278$ ($r_{\text{partial}} = -.209$). Removing RC did not significantly decrease explained variance of RAT scores, $\Delta R^2 = -.025$, $\Delta F(1, 31) = 1.419$, $p = .243$. The remaining predictors were significant: $b = 5.21$, $SE = 1.83$, $\beta = .423$, $t(32) = 2.84$, $p = .015$ (Bonferroni-corrected p value), for RR, and $b = -1.76$, $SE = .67$, $\beta = -.352$, $t(32) = -2.629$, $p = .025$ (Bonferroni-corrected p value), for associative fixed RT.

Discussion

Executive measures and their involvement in RAT

Our results suggest that producing associative and dissociative verbal responses reflects at least partially distinct processes. Functionally, the ACT associative RT (especially in fixed sequences) is analogous to HSCT response initiation and semantic fluency—that is, measures shown to predominately engage lexical–semantic functioning (Henry & Crawford, 2004; Shao, Janse, Visser, & Meyer, 2014; Whiteside et al., 2016). On the other hand, dissociative responses require additional inhibition of the dominant associative responses evoked by

Fig. 5 Associations between RAT performance and (A) associative fixed response time, (B) response remoteness, and (C) response commonness in the associative chain test

word cues and shifting to a semantically unrelated semantic cluster in memory. This additional processing substantially increased the average time required for producing words. As has been shown in previous research (Allen et al., 2008; Collette et al., 2001; de Zubizaray, Zelaya, Andrew, Williams, & Bullmore, 2000), response initiation (delivering associates) and response suppression (delivering dissociates) employ different brain circuits: Initiating an associative response has been coupled with activation of the inferior frontal gyrus and tempo-parietal areas implicated in the storage and retrieval of semantic information (see also Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), whereas response suppression has been associated with additional activation in the left dorsolateral prefrontal cortex and the orbitofrontal cortex—that is, areas involved in executive attention and working memory (Diamond, 2013). This neuroimaging evidence supports conceptualizing inhibition cost as a measure reflecting the executive attentional functions responsible for restraining access to strong but inappropriate associative responses and reducing goal-irrelevant information interference (Friedman & Miyake, 2004). Notably, this inhibition time cost was reliable across the conditions and tasks. Our data showed that delivering associative and dissociative responses in alternation also resulted in a reliable increase in response time. Importantly, the switching cost was approximately two-fold for dissociative responses as compared to associative responses (.71 and .36 s, respectively) and the inhibition cost was higher in alternating than in fixed chains (2.40 and 2.05 s, respectively), whereas both measures were positively correlated. Such a pattern indicates that inhibiting inappropriate prepotent responses and flexible switching among different rules and/or semantic clusters recruit common executive resources (i.e., imposing inhibition load inflates switching cost, and vice versa). This account is consistent with neurobiological and behavioral evidence showing that switching and inhibition engage overlapping prefrontal brain networks (Dove, Pollmann, Schubert, Wiggins, & Von Cramon, 2000; Ravizza & Carter, 2008) and processes, which resolve interference from the recently active goals, strategies, and/or response sets (Hyafil, Summerfield, & Koechlin, 2009).

Interestingly, neither inhibition nor switching was significantly associated with RAT performance, indicating that finding remote associates does not crucially involve these executive functions. On the other hand, RAT was negatively predicted by the latency of associative responses (in fixed sequences). This observation is in line with previous evidence for a positive link between RAT and associative fluency (Benedek, Könen, & Neubauer, 2012; Benedek & Neubauer, 2013; Marko & Riečanský, 2018; Mednick et al., 1964). However, in generating chains, each response redefines the applicability of the following ones, and therefore increased involvement of executive updating might be assumed. This account would be consistent with studies indicating that

RAT score weakly correlates with working memory capacity (Lee, Huggins, & Theriault, 2014; Lee & Theriault, 2013), a measure engaging executive updating (Miyake et al., 2000). However, RAT was not reliably correlated with latencies of associative responses in alternating sequences, which requires even a more substantial updating and shifts of the activated semantic sets. Taken together, our data indicate that the executive involvement in RAT is not substantial.

Lexical–semantic measures and their involvement in RAT

Association type had a robust impact on individuals' response remoteness and commonness. As expected, the remoteness of word pairs was much higher for associative than for dissociative responses. Dissociating was also coupled with more uncommon responses (lower commonness), which might be attributed to a general strategy of using less frequent words when retrieving unrelated word pairs. Interestingly, both commonness and remoteness covaried across all ACT conditions, suggesting that these measures may reflect reliable characteristics of the lexical–semantic system, in which individuals differ. Moreover, these characteristics were not associated with ACT inhibition and switching cost, indicating different underlying processes. Importantly, both commonness and remoteness were associated with RAT performance. In line with Mednick's (1962) model, these relations could be attributed to increased connectivity of lexical–semantic networks in creative individuals (i.e., flatter associative hierarchies as measured by RAT; see Kenett et al., 2014). Such cognitive feature would promote the ability to reach broader lexical–semantic and associative networks and thus increase the likelihood to access and retrieve more uncommon and remote associative elements. Alternatively, it has been suggested that the higher uncommonness of responses in creative individuals might be due to higher verbal fluency (i.e., creative individuals are able to quickly skim through the first common responses and thus to reach later responses, which are typically more uncommon and remote, sooner; see Benedek & Neubauer, 2013). Although this effect may apply when producing continuous associative responses to a fixed cue, it is implausible in case the cue word changes (as in the ACT). Consistent with this expectation, our data showed that the relation between RAT and the lexical–semantic measures held after RT in the associative fixed condition was controlled for (partial correlation coefficients $r = .550$, and $-.399$ for remoteness and commonness, respectively). These findings therefore suggest that the RAT predominantly reflects individuals' lexical–semantic network connectivity and semantic processing rather than executive functioning (i.e., inhibition, switching and/or verbal fluency, which is sometimes also considered an executive function; see Diamond, 2013). Importantly, it could be speculated that the relationship between RAT and the remoteness of associative and dissociative

responses may be conceptually distinct. On the one hand, the remoteness of free associative responses presumably emerges from the pattern of spontaneous activation within lexical–semantic networks (i.e., an automatic, bottom-up modulation). On the other hand, the remoteness of an unrelated (dissociative) response may be driven by the ability to suppress this automatic pattern of activation and to bias retrieval toward dissociated semantic clusters. In line with this distinction, it has been proposed that such a controlled retrieval mechanism is engaged when semantic information is not evoked automatically—that is, when lexical–semantic connections are weak (Badre & Wagner, 2007). Since bottom-up cues are not effective for retrieving suitable candidates during both the RAT and when generating dissociated words, the involvement of this semantic control mechanism may be required. Resolving of this conceptual consideration awaits further empirical research.

Conclusions

The present findings indicate that solving RAT problems engages a coherent cognitive ability that allows for reaching normatively remote elements of lexical–semantic and associative memory. A detailed analysis of associative behavior promotes the idea that this ability reflects individual differences in associative processing and structures rather than executive attentional functions (i.e., inhibition and switching). This does not mean that executive functions are not engaged in the RAT, but they presumably do not drive individual differences in accessing and combining remote associates among healthy adult individuals. Importantly, when assessing associative processes, the ACT may have several advantages over the RAT. In addition to response remoteness, ACT provides information about response fluency and commonness. Moreover, associative chains can be used to sample individuals’ “train of thought,” to which network theory analysis and computations can be applied. Perhaps the biggest advantage of this approach is its flexibility—using relatively simple rules, additional measures (e.g., executive) can be derived. Such measures may be valuable when studying associative processes in human thinking and creativity.

References

- Allen, P., Mechelli, A., Stephan, K. E., Day, F., Dalton, J., Williams, S., & McGuire, P. K. (2008). Fronto-temporal interactions during overt verbal initiation and suppression. *Journal of Cognitive Neuroscience*, 20, 1656–1669. <https://doi.org/10.1162/jocn.2008.20107>
- Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, 45, 2883–2901. <https://doi.org/10.1016/j.neuropsychologia.2007.06.015>
- Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., ... Neubauer, A. C. (2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage*, 90, 99–106. <https://doi.org/10.1016/j.neuroimage.2013.12.046>
- Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 273–281. <https://doi.org/10.1037/a0027059>
- Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick’s model on creativity-related differences in associative hierarchies: Evidence for a common path to uncommon thought. *Journal of Creative Behavior*, 47, 273–289. <https://doi.org/10.1002/jocb.35>
- Benedek, M., Panzner, L., Jauk, E., & Neubauer, A. C. (2017). Creativity on tap? Effects of alcohol intoxication on creative cognition. *Consciousness and Cognition*, 56, 128–134. <https://doi.org/10.1016/j.concog.2017.06.020>
- Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers*, 35, 634–639. <https://doi.org/10.3758/BF03195543>
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67, 380–400. <https://doi.org/10.1037/h0040373>
- Collette, F., Linden, M., Van Der Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2001). The Functional anatomy of inhibition processes investigated with the hayling task. 267, 258–267. <https://doi.org/10.1006/nimg.2001.0846>
- Coney, J., & Serna, P. (1995). Creative thinking from an information processing perspective: A new approach to Mednick’s theory of associative hierarchies. *Journal of Creative Behavior*, 29, 109–132. <https://doi.org/10.1002/j.2162-6057.1995.tb00740.x>
- Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingrover, H., Wetzels, R., Grasman, R. P. P., ... Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review*, 23, 640–647. <https://doi.org/10.3758/s13423-015-0913-5>
- Dailey, D. P. (1978). An analysis and evaluation of the internal validity of the remote associates test: What does it measure? *Educational and Psychological Measurement*, 38, 1031–1040. <https://doi.org/10.1177/001316447803800421>
- Davelaar, E. J. (2015). Semantic search in the remote associates test. *Topics in Cognitive Science*, 7, 494–512. <https://doi.org/10.1111/tops.12146>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & Von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: An event-related fMRI study. *Cognitive Brain Research*, 9, 103–109. [https://doi.org/10.1016/S0926-6410\(99\)00029-4](https://doi.org/10.1016/S0926-6410(99)00029-4)
- Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review*, 22, 13–37. <https://doi.org/10.3758/s13423-014-0665-7>
- Eysenck, H. J. (1995). *Genius: The natural history of creativity*. Cambridge, UK: Cambridge University Press.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, 133, 101–135. <https://doi.org/10.1037/0096-3445.133.1.101>
- Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, 18, 284–295. <https://doi.org/10.1037/0894-4105.18.2.284>
- Hyafil, A., Summerfield, C., & Koechlin, E. (2009). Two mechanisms for task switching in the prefrontal cortex. *Journal of Neuroscience*, 29, 5135–5142. <https://doi.org/10.1523/JNEUROSCI.2828-08.2009>

- Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 8, 407:1–16. <https://doi.org/10.3389/fnhum.2014.00407>
- Kenett, Y. N., 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*, (January), 201717362. <https://doi.org/10.1073/pnas.1717362115>
- Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology*, 65, 71–93. <https://doi.org/10.1146/annurev-psych-010213-115154>
- Lee, C. S., Huggins, A. C., & Theriault, 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, 446–460. <https://doi.org/10.1037/a0036773>
- Lee, C. S., & Theriault, 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>
- Marko, M., & Riečanský, I. (2018). Sympathetic arousal, but not disturbed executive functioning, mediates the impairment of cognitive flexibility under stress. *Cognition*, 174, 94–102. <https://doi.org/10.1016/j.cognition.2018.02.004>
- Marupaka, N., Iyer, L. R., & Minai, A. A. (2012). Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. *Neural Networks*, 32, 147–158. <https://doi.org/10.1016/j.neunet.2012.02.004>
- Mednick, M. T., Mednick, S. A., & Jung, C. C. (1964). Continual association as a function of level of creativity and type of verbal stimulus. *Journal of Abnormal Psychology*, 69, 511–515. <https://doi.org/10.1037/h0041086>
- Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, 69, 220–232. <https://doi.org/10.1037/h0048850>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “{Frontal} {Lobe}” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, 46, 2924–2935. <https://doi.org/10.1016/j.neuropsychologia.2008.06.006>
- Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? {Predictors} of verbal fluency performance in older adults. *Frontiers in Psychology*, 5, 772. <https://doi.org/10.3389/fpsyg.2014.00772>
- Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the Remote Associates Test. *Cognition*, 128, 64–75. <https://doi.org/10.1016/j.cognition.2013.03.001>
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, 94, 14792–14797. <https://doi.org/10.1073/pnas.94.26.14792>
- Weiland, H., Bambini, V., & Schumacher, P. B. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, 8, 1–17. <https://doi.org/10.3389/fnhum.2014.00583>
- White, H. A., & Shah, P. (2006). Uninhibited imaginations: Creativity in adults with attention-deficit/hyperactivity disorder. *Personality and Individual Differences*, 40, 1121–1131. <https://doi.org/10.1016/j.paid.2005.11.007>
- Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, 23, 29–34. <https://doi.org/10.1080/23279095.2015.1004574>
- Worthen, B. R., & Clark, P. M. (1971). Toward an improved measure of remote associational ability. *Journal of Educational Measurement*, 8, 113–123.
- Wu, C., Zhong, S., & Chen, H. (2016). Discriminating the difference between remote and close association with relation to white-matter structural connectivity. *PLoS ONE*, 11, e165053:1–16. <https://doi.org/10.1371/journal.pone.0165053>
- Wu, C. L., & Chen, H. C. (2017). Normative data for Chinese compound remote associate problems. *Behavior Research Methods*, 49, 2163–2172. <https://doi.org/10.3758/s13428-016-0849-3>
- Xiao, X., Zhao, D., Zhang, Q., & Guo, C. Y. (2012). Retrieval of concrete words involves more contextual information than abstract words: Multiple components for the concreteness effect. *Brain and Language*, 120, 251–258. <https://doi.org/10.1016/j.bandl.2011.09.006>
- de Zubizaray, G. I., Zelaya, F. O., Andrew, C., Williams, S. C. R., & Bullmore, E. T. (2000). Cerebral regions associated with verbal response initiation, suppression and strategy use. *Neuropsychologia*, 38, 1292–1304.
- Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, 6, 21–29. <https://doi.org/10.1107/S0907444909031205>