Development and Preliminary Exploration of a Chinese Word Remote Association Test

Authors: Chen Jing, Xu Tianwei, Qiu Jiang, Chen Qunlin, Chen Qunlin

Date: 2023-09-10T14:36:52+00:00

Abstract

Remote Associates Test (RAT) and its variant Compound Remote Associates Test (CRAT) are the most commonly used experimental materials for revealing the cognitive neural mechanisms of convergent thinking and insight, and are also standard assessments for evaluating individual creative potential. Compared to CRAT, RAT is more suitable for cross-cultural research; however, there is currently a lack of Chinese remote associates materials equivalent to RAT. Based on the classic associative hierarchy theory and connection characteristics within the Chinese linguistic context, this study developed 80 sets of Chinese Word Remote Associates Test (CWRAT); additionally, a graded scoring model was established. Drawing upon latent class analysis methodology, a latent class analysis was conducted on the 80 items using connection type as the classification index, and the impact of classification results on various item indicators (e.g., insight experience) was investigated. Validity tests demonstrated that CWRAT is significantly correlated with creative analogy, creative achievement, and openness. According to the statistical indicators of latent class analysis, CWRAT can be divided into three categories; it was found that categories C1 and C3, characterized by high symbolic/metaphorical probability connection types, exhibit higher insight experience than category C2, which is characterized by high characteristic probability connection types, while category C2 shows higher pass rates and lower reaction times than category C3. This measurement enriches the existing repository of Chinese creative assessment tools and provides new experimental materials for cross-cultural comparisons of creativity.

Full Text

Development and Preliminary Exploration of the Chinese Word Remote Associates Test

Jing Chen[1], Tianwei Xu[2],[3], Jiang Qiu[1], Qunlin Chen[1]

[1] Faculty of Psychology, Southwest University; Key Laboratory of Cognition and Personality, Ministry of Education, Chongqing, 400715, China
[2] College of Science, Qiongtai Normal University, Haikou, 571127, China
[3] Key Laboratory of Child Cognition & Behavior Development of Hainan Province, Haikou, 571127, China

Abstract

The Remote Associates Test (RAT) and its variant, the Compound Remote Associates Test (CRAT), are the most commonly used experimental materials for revealing the cognitive neural mechanisms of convergent thinking and insight, as well as for assessing individual creative potential. Compared with CRAT, RAT is more suitable for cross-cultural research; however, there is currently a lack of Chinese remote associates materials equivalent to RAT. Based on classical associative hierarchy theory and the associative characteristics in the Chinese linguistic context, this study developed 80 items of the Chinese Word Remote Associates Test (CWRAT). Additionally, a graded scoring model was established. Using latent class analysis methodology and taking associative methods as classification indicators, a latent class analysis was conducted on the 80 items to explore the impact of classification results on various item indicators (e.g., sense of insight). Validity tests showed that CWRAT was significantly correlated with creative analogy, creative achievement, and openness. According to the statistical indicators from the latent class analysis, CWRAT items could be divided into three categories. It was found that categories C1 and C3, characterized by a high probability of symbolic/metaphorical associative methods, exhibited higher insight than category C2, which was characterized by a high probability of feature-based methods. Moreover, category C2 showed higher pass rates and lower reaction times than category C3. This measure enriches the existing repository of Chinese creativity assessment tools and provides new experimental materials for cross-cultural comparison of creativity.

Keywords: Remote Associates Test; associative hierarchy; creativity; semantic distance

1.1 The Development of Remote Associates Test

Mednick (1962) proposed that creative idea generation depends on spontaneous semantic associative processing, which involves connecting elements or concepts from memory in novel ways that meet specific requirements. From an individual differences perspective, the ability to establish connections between seemingly unrelated remote concepts represents one manifestation of creative potential. Empirical studies have found that highly creative individuals possess more flexible semantic networks in free association tasks, whereas less creative individuals exhibit relatively rigid networks (He et al., 2020; Kenett et al., 2016). Based on associative hierarchy theory, Mednick (1967) developed the Remote Associates Test (RAT) as a classic task for measuring creativity, which has been widely used

in research on the mechanisms of convergent thinking and insight and has been applied across different countries and languages (Behrens & Olteţeanu, 2020), making it suitable for cross-cultural comparisons. However, due to the structural particularities of the Chinese language (Hansen, 1993), the RAT task has not been widely used domestically, partly because a standardized Chinese version has not been developed. Therefore, this study will follow standard test development procedures to create a Chinese Word Remote Associates Test (CWRAT) suitable for the Chinese linguistic context.

1.2 Development and Differences of Chinese RAT

Classic RAT items consist of three unrelated or weakly related cues, requiring participants to find a single answer that can be associated with each of the three cues separately. The associative methods include synonyms, semantic connections, or forming compound words. Many researchers have revised or redeveloped RAT according to the characteristics of their study populations or linguistic contexts. Bowden and Jung-Beeman (2003) argued that the classic RAT's complex associative methods and high difficulty hindered large-scale test development and administration. Consequently, they developed the Compound Remote Associates Test (CRAT) to meet the needs of cognitive neuroscience research. CRAT employs only one associative method: forming compound words. CRAT items are easy to develop and possess insight-like characteristics.

Due to the particularities of the Chinese linguistic context, the existing English versions of RAT cannot be directly applied to Chinese participants through simple translation or revision. After identifying differences in linguistic rules and features between Chinese and English, Jen et al. (2004) developed the Chinese Compound Remote Associates Test (CCRAT) based on the same principle as CRAT, with only one associative method: forming character pairs. Huang et al. (2012) constructed the Chinese Word Remote Associates Test (CWRAT) based on the principle of classic RAT, which includes three associative methods. Currently, only a few researchers have expanded or explored CWRAT (Huang et al., 2012; Xiao et al., 2016). A comparison of different RAT versions is presented in Table 1 .

Table 1: Comparison of content across different RAT versions

Version	Examples	Connection Type
CCRAT	Same-fire-head; Flower-Friend-Scout; 板-洞-色	Form compound words
CWRAT	市场-结束-日落	Form two-character words
Match	Flower girl-Girlfriend-Girl scout; 黑板-黑洞-黑色	

1.3 Comparison of Chinese and English RAT

Due to linguistic and cultural differences, various versions of RAT have emerged. These version differences lead to several problems: (1) different language diffi-

culties, where simple translation creates test heterogeneity; (2) even when translation is possible, differences in the number and strength of associations between words in different languages still cause heterogeneity; and (3) each test developer has different background knowledge and cognitive preference styles, which may result in non-standardized test creation methods (Behrens & Olteţeanu, 2020).

CCRAT was adapted from CRAT and has developed a large item bank due to its ease of construction (Shen et al., 2016). However, the smallest meaningful units differ between Chinese and English, and CCRAT focuses more on word construction tasks with less involvement of semantic or creative associations. Therefore, CCRAT uses a "character-to-character pair formation" approach instead of CRAT's "word-to-word compound formation," achieving only formal similarity while differing in the essence of association from the original CRAT. Some tests have not adequately considered "remoteness" as a primary factor determining item difficulty during development (Marko et al., 2019). For example, non-standard construction methods result in target words that are all high-frequency vocabulary (Wu et al., 2017), and excessive experience with high-frequency stimuli may actually hinder creativity (Gupta et al., 2012). Moreover, due to the particularities of the Chinese language, participants' cultural level, linguistic thinking ability, and vocabulary size also affect test performance, presenting confounding variables (Wu et al., 2017; Xiao et al., 2016). Therefore, despite its rapid development, CCRAT still has numerous deficiencies in the Chinese context.

CWRAT is difficult to develop due to the diversity of cue word associations and non-unique answers. The multiple associative methods in CWRAT cause item heterogeneity, leading researchers to question whether different cognitive systems or associative mechanisms are required to process different connection types (Marko et al., 2019). Furthermore, the classification of associative methods is too general; for instance, the "semantic connection" method could be further subdivided into "conceptual," "symbolic/metaphorical," and other levels. Among these, metaphorical connection (imagery) is an important factor in item difficulty, requiring individuals to have higher flexible integration of knowledge and semantic representations (Marko et al., 2019). In reality, existing CWRAT has not considered important parameters such as item heterogeneity and imagery. Additionally, the limited number of CWRAT items reduces usability and data reliability (Bowden & Jung-Beeman, 2003; Shen et al., 2016). Moreover, in some items, the semantic activation spaces among cues overlap extensively, and requiring a single answer may cause other correct or more creative answers to be overlooked. In summary, CWRAT suffers from issues such as lack of standardization and incomplete evaluation indicators, requiring further improvement.

2 Problem Statement

In summary, CCRAT has a simple format and numerous items, but it resembles a word construction task of "forming words from characters." While it appears consistent with the logic of remote association, it primarily assesses individual differences in semantic knowledge accumulation and does not align with the creative definition of novel combination and connection of concepts based on problem goals.

In contrast, although CWRAT is more difficult to develop, it better conforms to associative hierarchy theory and is homogeneous with classic RAT, making it suitable for cross-cultural creativity comparisons. Moreover, CWRAT is based on real-life conceptual connections, giving it high ecological validity. However, existing CWRAT suffers from limited item numbers, lack of reliability and validity testing, and non-standardized scoring criteria, which restricts its development. This study aims to develop a CWRAT suitable for the Chinese context based on associative hierarchy theory and following standardized test development procedures.

The CWRAT developed in this study follows the same principles as classic RAT (1962) and CWRAT (2012), with multiple associative methods between cues and answers. Synthesizing previous research, the three associative methods were further subdivided into twelve types (Marko et al., 2019; Mednick, 1962; Zhao, 2014) to explore the influence of associative methods on creative problem solving. Some studies suggest that RAT or CWRAT overly relies on language ability and vocabulary; therefore, this study's item design emphasizes imagistic thinking to reduce language influence (Xiao et al., 2016). Additionally, this study does not require unique answers but instead adopts a graded scoring system for answer evaluation.

3.1 Participants

A total of 108 university students were randomly recruited to participate in the development of CWRAT materials (mean age = 19.99, age range = 17-25 years, 11 males). Participants were randomly divided into two groups for online (Wenjuanxing) and laboratory data collection. Among them, 58 participants completed the online experiment, while the other 50 completed data collection in the laboratory. The experimental tasks were identical in both conditions, with the only difference being that the laboratory task used E-Prime software to record reaction time measures. E-Prime 2.0 was used to program the experiment and record participants' reaction times and responses. Before the experiment, participants were familiarized with the procedure and signed informed consent forms. All participants received compensation after completing the experiment.

3.2.1 Initial Development and Screening of CWRAT

Based on RAT principles, 279 items were created using words selected from the internet, daily life, and the *Chinese Language Dictionary*. Each item consisted of three cues and one reference answer.

Five researchers in the field of creativity were invited to conduct preliminary ratings on a series of item indicators. First, viewing the three cues and reference answer as an integrated whole, they rated each item's interestingness, match quality, sense of insight, and difficulty on a 5-point scale. Second, they evaluated the associative distance and connection type between individual cues and the answer. The connection types included feature/attribute relations, composition/inclusion relations, functional relations, spatial/temporal relations, categorical relations, imagery relations, synonym relations, antonym relations, symbolic/metaphorical relations, synthetic relations, phonological relations, and textual meaning relations.

After obtaining ratings from the five raters on all test items, 80 items were selected based on insight, semantic distance, interestingness, difficulty, and match quality (see supplementary materials). Some items were modified to form the final test, which underwent a new round of rating and compilation for interestingness, insight, match quality, difficulty, and connection types (see supplementary materials), using the same rating procedure described above.

3.2.2 Experimental Procedure

In both online and laboratory experiments, participants first completed 80 CWRAT tasks (see Figure 1 [Figure 1: see original paper]). They were required to think of an answer that could be associated with each of the three given cues simultaneously within a 30-second time limit per item. After completing the CWRAT, participants were administered the Creative Achievement Questionnaire, the NEO Five-Factor Inventory (NEO-FFI), and the Analogical Reasoning Test.

3.2.3 Scoring Criteria

Four researchers in the field of creativity were invited to grade participants' responses to CWRAT using a graded scoring system: 0 points indicated a completely inadequate answer that failed to connect with all three cues; 2 points indicated the reference answer or answers judged adequate based on objective knowledge; 1 point indicated an answer that was not correct but was related to the 2-point answer, showing correct thinking direction but still distant from the ideal answer (e.g., if the reference answer was "pipa" [a Chinese lute], a participant's answer of "still holding the pipa half-hidden behind the face" [a line from a poem] would receive 1 point) (see supplementary materials for answer scoring and alternative answers). First, the mean score from the four expert raters was calculated as the item score, and the CWRAT total score was obtained by

summing scores across all items. Additionally, the pass rate for each item was calculated as the sum of all participants' scores for that item divided by 160 (the maximum possible score across 80 items). The inter-rater reliability (ICC) was 0.947.

3.3.1 Creative Achievement Questionnaire (CAQ)

The Creative Achievement Questionnaire (Chen, 2014; Carson et al., 2005) is a domain-specific creativity measurement tool covering ten fields generally considered to require high creativity: visual arts, music, creative writing, dance, architectural design, humor, inventions, scientific discovery, drama and film, and culinary arts. Creative achievement scores typically show a positively skewed distribution, indicating that only a minority of individuals demonstrate high creative achievement. Research shows that CAQ has relatively high reliability and validity compared to similar measures, with test-retest reliability of 0.81 and internal consistency reliability of 0.96, and moderate correlations with divergent thinking and openness personality (Chen, 2014).

3.3.2 NEO Five-Factor Inventory (NEO-FFI)

The NEO Five-Factor Inventory (NEO-FFI) includes five dimensions: Extroversion, Neuroticism, Openness to Experience, Agreeableness, and Conscientiousness, comprising 60 items (Costa & McCrae, 2010; J. Li et al., 2018). Openness has commonly been used as a criterion in the development of creativity measurement tools, and this study only considered the Openness dimension in its analyses.

3.3.3 Analogical Reasoning Test (ART)

Analogical reasoning is widely considered a crucial cognitive process for creativity. Research shows that the semantic distance in cross-domain analogical reasoning is significantly positively correlated with the originality of divergent thinking, suggesting that cross-semantic domain analogical reasoning is a key mechanism of creativity (Zeng, 2021). ART contains 78 items divided into four conditions: cross-domain analogical reasoning, within-domain analogical reasoning, cross-domain false analogical reasoning, and within-domain false analogical reasoning. Correct answers receive 1 point, while incorrect answers receive 0 points.

3.4 Data Analysis

SPSS 26.0 and Mplus were used for statistical analysis. Correlation analysis was conducted between participants' total CWRAT scores and CAQ, Openness, and ART scores to test criterion validity. Each item score was correlated with the three criterion test scores, and the correlation coefficients were centered and summed to obtain the item's standardized correlation coefficient z(r).

Reaction times, pass rates, and standardized correlation coefficients z(r) for individual items were compiled (see supplementary materials). Latent class analysis (Muthén & Muthén, 2010) was employed to classify the 80 items using associative methods as classification indicators. The modified BCH method was used to examine the effects of classification results (predictor variables) on item subjective difficulty, interestingness, match quality, insight, pass rate, and reaction time (outcome variables).

4 Results

Descriptive statistics for demographic variables and measurement scores are presented in Table 2 , and descriptive statistics for item difficulty, interestingness, and other indicators are shown in Table 3 .

To exclude potential systematic differences between data sources (Wenjuanxing vs. laboratory), independent samples t-tests were conducted on common measurement indicators (CWRAT, CAQ, ART, and Openness scores) from different sources. Results showed no significant differences, indicating data homogeneity across sources.

Correlation analysis revealed that participants' total CWRAT scores were significantly positively correlated with CAQ scores ($r = .22$, $p = 0.024$), Openness scores ($r = .24$, $p = 0.012$), and ART scores ($r = .48$, $p < 0.001$) (see Figure 2 [Figure 2: see original paper]).

Latent class models with 2-5 classes were extracted, and fit indices are summarized in Table 4 . As shown in Table 4, fit indices AIC, BIC, and aBIC increased monotonically with the number of classes, and the Entropy value was highest for the two-class solution, indicating high classification accuracy for the two-class model. However, the BLRT p-value for the three-class model reached significance, suggesting that the three-class model was significantly better than the two-class model. Research indicates that BLRT is the most effective index for determining the number of latent classes (Nylund et al., 2007).

Table 2: Descriptive statistics of participants' age and creativity measurement scores

Variable	Min	Max	M	SD
Creative Achievement				
Creative Analogy				
Openness Personality				
CWRAT Total Score				

Table 3: Descriptive statistics of item difficulty, interestingness, and other indicators

Variable	Min	Max	M	SD
z(r) Total Score				

Based on the fit indices and theoretical considerations, a three-class solution was deemed appropriate. The distribution of conditional probability sums for the three cue words across the 12 connection types for the three classes (C1, C2, C3) is presented in Figure 3 [Figure 3: see original paper]. The results show that the primary characteristic of Class C1 is a high probability of symbolic/metaphorical connections, Class C2 shows high probability of feature-based connections, while Class C3 shows high probabilities in composition/inclusion relations, functional relations, and symbolic/metaphorical relations.

Results from the regression mixture model are presented in Table 5 . The findings indicate that the insight level of Class C3 was significantly higher than that of Class C2, $\Delta\ ^2(1) = 4.325$, $p = 0.038$. The pass rate of Class C1 was significantly higher than that of Class C3, $\Delta\ ^2(1) = 7.635$, $p = 0.006$, and the pass rate of Class C2 was significantly higher than that of Class C3, $\Delta\ ^2(1) = 19.853$, $p < 0.001$. The reaction time of Class C3 was significantly longer than that of Class C1, $\Delta\ ^2(1) = 6.197$, $p = 0.013$, and significantly longer than that of Class C2, $\Delta\ ^2(1) = 15.525$, $p < 0.001$. No significant differences were found among the three latent classes in terms of item difficulty, interestingness, or match quality.

Table 4: Comparison of fit indices for latent class analysis

Class	BLRT	Entropy
2		0.4/0.6
3		0.3875/0.46250/0.15000
4		0.47500/0.25000/0.10000/0.17500
5		0.225/0.08750/0.36250/0.13750

Table 5: Mean differences in item indicators across the three item categories

Indicator	Class 1: M (SE)	Class 2: M (SE)	Class 3: M (SE)	$^2(2)$
Insight	2.539 (0.085)	3.308 (0.116)	4.678 (0.065)	21.315**
Pass Rate	2.538 (0.138)ab	0.526 (0.046)a	2.642 (0.106)	
Reaction Time	3.448 (0.106)	4.754 (0.055)	2.484 (0.097)a	
Difficulty	0.593 (0.040)a	2.485 (0.136)	3.454 (0.190)	
Interestingness	4.599 (0.139)	2.854 (0.150)b	0.360 (0.033)b	
Match Quality				15.770**

Note: Groups with different letter subscripts (a, b, c) differ significantly at $p < 0.05$. * represents $p < .05$; ** represents $p < .001$.

Figure 2: Correlation analysis between CWRAT scores and CAQ, Openness, and ART scores

Figure 3: Distribution of conditional probability sums for three cue words across 12 connection types in items

Note: 1 = Feature/attribute relation, 2 = Composition/inclusion relation, 3 = Functional relation, 4 = Spatial/temporal relation, 5 = Categorical relation, 6 = Imagery relation, 7 = Synonym relation, 8 = Antonym relation, 9 = Symbolic/metaphorical relation, 10 = Synthetic relation, 11 = Phonological relation, 12 = Textual meaning relation.

5 Discussion

This study aimed to develop a remote associates task suitable for Chinese populations and to classify the potential associative methods between cues and answers into twelve categories. Additionally, it modified the traditional correct-answer scoring to a graded scoring system using reference answers. Criterion-related validity analysis revealed that CWRAT scores were significantly positively correlated with CAQ, Openness, and ART scores. Latent class analysis showed that CWRAT items could be divided into three categories, with significant differences among the three latent classes in insight, pass rate, and reaction time.

Numerous studies have shown that Openness to Experience can significantly predict individual creative thinking ability (Zare & Flinchbaugh, 2019), and Openness is considered a gold standard criterion for validating creativity measures (Feist, 2018). From a structural perspective of associative task types, creative problem-solving processes like RAT require integrating bottom-up free association with top-down cognitive control, potentially benefiting from divergent thinking and thus depending partly on flexibility (Olteţeanu & Falomir, 2015; Zhang et al., 2020). Analogical reasoning constitutes an important component of creative thinking (Dunbar & Blanchette, 2001), with creative analogical reasoning emphasizing mapping relationships between concepts during problem solving. Green et al. proposed that long-distance analogical mapping facilitates creative thinking and represents a core process in establishing creative connections across remote concepts. The present findings suggest that remote associative thinking may involve similar cognitive processes, requiring identification of key links between semantically distant words to generate novel connections.

Latent class analysis was used to divide items into three categories, and regression mixture modeling explored the effects of these latent classes on other item indicators. Regarding insight, Class C2 was significantly lower than Class C3 and also lower than Class C1. The most distinctive feature of Classes C1 and C3 compared to C2 is their higher probability of symbolic/metaphorical relations, whereas Class C2 shows higher probability of feature-based relations. Both pass rate and reaction time are indicators of item difficulty. The results showed that Class C2 had higher pass rates and lower reaction times than Class C3, indicating that Class C2 items were less difficult. Compared to novel uses

of concepts or objects, features represent conventional associations to concepts and are believed to involve less creative cognitive demand (Fink et al., 2010). Insight is defined as the connection and recognition of remote concepts, and the experience of insight often follows the breaking of mental sets and transformation of problem representations (DeYoung et al., 2008), whereas conventional feature-based connections may involve less expansion of semantic space. Symbolism/metaphor is a rhetorical device involving abstract connections between similar attributes of two concepts in semantic memory and is considered an important indicator for measuring creative thinking (Leung et al., 2012). Furthermore, analogical reasoning is highly correlated with metaphor, with novel metaphors understood through analogy. Analogical reasoning helps integrate common or similar features of concepts from different semantic domains and map their relationships and structures; for example, Chinese scientists improved spacesuits through analogy with crayfish (Yang et al., 2018). In the CWRAT developed in this study, there are numerous symbolic/metaphorical connections, such as "bird-airplane" and "peony-wealth." Therefore, Class C2 items, characterized by high-probability feature-based relations, may tend toward conventional semantic associations involving non-insightful or low-insight solutions. In contrast, Class C1 and C3 items, containing higher-probability symbolic/metaphorical connections, possess more characteristics of creative, insightful solutions.

Currently, classic RAT and its extended versions have been widely applied in psychological measurement. This study provides a sufficient number of Chinese word RAT items to help researchers explore individual differences in the neural mechanisms underlying convergent thinking and insight problem solving, and to assist educators in understanding students' creative potential in educational assessment contexts, thereby scientifically cultivating students' creative personalities. Additionally, RAT can be used in research on psychopathology, emotion, success and failure experiences, and other areas to reveal underlying mechanisms. In summary, this study developed and selected 80 final items suitable for use in cognitive neuroscience and education fields, incorporating various normative indicators to help other researchers select appropriate items according to their experimental purposes and populations.

This study has several limitations. For instance, we recruited raters to grade answers using a graded system, but this depends on raters' subjective understanding of items. Future research should adopt more objective methods for answer evaluation, such as calculating semantic distance between cues and answers based on corpora.

References

Zeng, R. (2021). *The influence of semantic network structure on remote analogical reasoning and its brain functional basis* [Master's thesis, Southwest University].

Chen, Q. (2014). *The relationship between creative achievement and cognitive*

flexibility [Master's thesis, Southwest University].

Xiao, W., Yao, X., & Qiu, Y. (2016). Developing a Chinese version of the Remote Associates Test (RAT) using item response theory. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2.

Zhao, J. (2014). *The influence of "student-based teaching" on primary school students' creative thinking* [Master's thesis, Nanjing Normal University].

Behrens, J. P., & Olteţeanu, A.-M. (2020). Are All Remote Associates Tests Equal? An Overview of the Remote Associates Test in Different Languages. *Frontiers in Psychology*, 11.

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

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, 17(1), 37-50.

Costa, P. T., Jr., & McCrae, R. R. (2010). Bridging the gap with the five-factor model.

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in Neurosciences*, 38(9), 571–578.

DeYoung, C. G., Flanders, J. L., & Peterson, J. B. (2008). Cognitive abilities involved in insight problem solving: An individual differences model. *Creativity Research Journal*, 20(3), 278-290.

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: The case of analogy. *Trends in Cognitive Sciences*, 5(8), 334-339.

Feist, G. J. (2018). In search of the Creative Personality. *The Nature of Human Creativity*, 63-76.

Fink, A., Grabner, R. H., Gebauer, D., Reishofer, G., Koschutnig, K., & Ebner, F. (2010). Enhancing creativity by means of cognitive stimulation: Evidence from an fMRI study. *NeuroImage*, 52(4),

Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, 20(1), 70-76.

Gupta, N., Jang, Y., Mednick, S. C., & Huber, D. E. (2012). The road not taken: Creative solutions require avoidance of high-frequency responses. *Psychological Science*, 23(3), 288-294.

Hansen, C. J. T. J. o. A. S. (1993). Chinese ideographs and Western ideas. 52(2), 373-399.

He, L., Kenett, Y. N., Zhuang, K., Liu, C., Zeng, R., Yan, T., . . . Reasoning. (2020). The relation between semantic memory structure, associative abilities, and verbal and figural creativity. 27(2),

Huang, P., Chen, H., & Liu, C. (2012). The development of Chinese word remote associates test for college students. *Psychological Testing*, 59(4), 581-607.

Jen, C., Chen, H., Lien, H., & Cho, S. (2004). The development of the Chinese remote association test. *Research in Applied Psychology*, 21, 195-217.

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

Leung, A. K.-y., Kim, S., Polman, E., Ong, L. S., Qiu, L., Goncalo, J. A., & Sanchez-Burks, J. (2012). Embodied metaphors and creative "acts". *Psychological Science*, 23(5), 502-509.

Li, J., Zhao, Y., Kong, F., Du, S., Yang, S., & Wang, S. (2018). Psychometric assessment of the short grit scale among Chinese adolescents. *Journal of Psychoeducational Assessment*, 36(3), 291–296.

Marko, M., Michalko, D., & Riečanský, I. (2019). Remote associates test: An empirical proof of concept. *Behavior Research Methods*, 51(6), 2700-2711.

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

Mednick, S. A., & Mednick, M. T. S. (1967). *Examiner's manual: Remote associates test*. Houghton Mifflin.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide: Statistical analysis with latent variables: User's guide*. Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 535-569.

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

Shen, W., Yuan, Y., Liu, C., Yi, B., & Dou, K. (2016). The development and validity of a Chinese version of the compound remote associates test. *American Journal of Psychology*, 129(3), 245-258.

Wu, C.-L., Chang, Y.-L., & Chen, H.-C. (2017). Enhancing the measurement of remote associative ability: A new approach to designing the Chinese Remote Associates Test. *Thinking Skills and Creativity*, 24, 29-38.

Yang, W., Jin, Y., Qiu, J., & Zhang, Q. J. A. P. S. (2018). The effect of prototype difficulty and semantic similarity on the prototype activation. 50(3), 260.

Zare, M., & Flinchbaugh, C. J. H. P. (2019). Voice, creativity, and big five personality traits: A meta-analysis. 32(1), 30-51.

Zhang, W., Sjoerds, Z., & Hommel, B. (2020). Metacontrol of human creativity: The neurocognitive mechanisms of convergent and divergent thinking. *NeuroImage*, https://doi.org/https://doi.org/10.1016/j.neuroimage.2020.116572

Note: Figure translations are in progress. See original paper for figures.