Journal of Experimental Psychology: General

# The Workload Capacity of Semantic Search in Convergent Thinking

Linlin Shang, Daniel R. Little, Margaret E. Webb, Ami Eidels, and Cheng-Ta Yang

Online First Publication, September 9, 2021. http://dx.doi.org/10.1037/xge0001045

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General


ISSN: 0096-3445

https://doi.org/10.1037/xge0001045

# The Workload Capacity of Semantic Search in Convergent Thinking

Linlin Shang[1], Daniel R. Little[2], Margaret E. Webb[2], Ami Eidels[3], and Cheng-Ta Yang[1, 4]

[1] Department of Psychology, National Cheng Kung University

[2] Melbourne School of Psychological Sciences, The University of Melbourne

[3] School of Psychology, The University of Newcastle

[4] Institute of Allied Health Sciences, National Cheng Kung University

The present study used *Systems Factorial Technology* (Townsend & Nozawa, 1995) to investigate how people combine dual cues in semantic memory search. Our aims were (a) to understand how cues interact during the process of semantic search in convergent thinking and (b) to determine how workload capacity (i.e. cue-processing efficiency) is related to search performance. In two experiments, participants completed a typical convergent thinking test and a word production task. The results revealed that: (a) collective evidence supports similar patterns in cue-combination strategy despite individual differences in workload capacity, and (b) there exists a negative correlation between workload capacity and performance on convergent thinking test. A potential explanation is that, for the creative individual, loading many candidate answers leads to consumption of substantial processing resources that obtains as low workload capacity but also allows creative individuals to switch more easily from one candidate to another so that there is a higher probability of successfully producing an answer within a limited time. Our results further imply that workload capacity is a significant factor for the semantic search process in convergent thinking and provides new insight on the model of semantic search and creativity.

*Keywords:* convergent thinking, semantic search, workload capacity, systems factorial technology



The ability to search for a specific piece of information based on associative cues supports a number of important cognitive functions such as convergent thinking. Convergent thinking is a cognitive process by which information is assessed for fitness in a given situation (e.g., whether a brick could be used as a wine holder; Vries & Lubart, 2019). Convergent thinking tasks often require exhaustive evaluation of all of the relevant cues. For instance, a friend asking whether anyone remembers the name of their high school geography teacher provides both temporal ("high school") and nominal cues ("geography teacher"). Several answers can be found through either cue alone; however, only a few answers (potentially only one) will satisfy the constraint of being jointly associated with both cues simultaneously. Hence, cue-combination may play an important role in the process of semantic search through associative cues.

Previous studies suggested that semantic search is a dynamic process, yet details of the process remain unclear. During semantic search processes, cues first activate corresponding nodes stored in long-term memory (LTM) to form one or more search patches. These patches consist of all the candidate answers related to the cues. Then a sequential search is employed to check the candidate answers one at a time from local patches (e.g., relatively close associates such as subcategory or semantic similarity) to global patches (remote associates) until the final answer is determined (Davelaar, 2015; Hills et al., 2012; Hills et al., 2015). However, most studies (e.g., Smith et al., 2013) are concerned with the relationship between the candidate answers and how the candidate answers influence each other; this influence can interfere with the final result of the search. Hence, the details of how cues interact with each other are unknown. The present study switches the focus from the candidate answers to the cues and investigates whether the cues' relationship is related to the final search result.

The primary debate focuses on whether multiple cues interact with each other and, if they do interact, in what manner. Some theories assume that the initial search process (e.g., the process of spreading activation; Collins & Loftus, 1975) proceeds in parallel with no interaction between the cues during the search process. These theories assume that each cue, represented as a probabilistically stored vector of features, is matched in parallel against

Cheng-Ta Yang https://orcid.org/0000-0002-6561-763X

Linlin Shang is now at Donders Institute for Brain, Cognition and Behaviour, Radboud University.

This work was supported by grants from China Scholarship Council to Linlin Shang, Ministry of Science and Technology, Taiwan (MOST 105-2410-H-006-020-MY2, MOST 107-2410-H-006-055-MY2, 109-2410-H-006 -049 -MY3, and 109-2511-H-006 -002 -MY2 to Cheng-Ta Yang) and Australian Research Council (Discovery Project Grant DP160102360 to Daniel R. Little and Ami Eidels).

The results of the present study were reported at the annual meeting of Society for Mathematical Psychology, a virtual conference held from July 20th to 31st, 2020 (the unlisted link is https://youtu.be/JIU-biuUtVo). The experimental materials and raw data of this study can be downloaded at https://github.com/shangll/OData-ConvergThink.

Correspondence concerning this article should be addressed to Cheng-Ta Yang, Department of Psychology, National Cheng Kung University, No. 1, University Road, Tainan 701, Taiwan. Email: yangct@mail.ncku.edu.tw

memory traces (Flexser & Tulving, 1978); and the search terminates when any cue matches strongly enough (i.e. a self-terminating rule). However, most findings have demonstrated that multiple cues can produce a better search result than single cues because a combination of cues increases the probability of successful retrieval (Massaro et al., 1991). This result suggests that search is not based on a single successful cue. Massaro et al., (1991) suggest cues could be multiplicatively combined in an *enhancing integration* fashion (i.e. using their fuzzy logical model of perception; FLMP). According to the enhancing integration model, associations between feature vectors of cues and candidate solutions are activated at overall levels, increasing the joint probability over the union or even the sum of single-cue probabilities. This may imply a facilitatory relationship between cues.

Alternatively, some studies argue for a *weighted averaging* model (WAM; Anderson, 1974; Lederman et al., 1986) where multiple cues are assumed to operate in a nonlinear weighted fashion (Wiles et al., 1991). That is, one of the cues finally dominates via a compromising process, thus suggesting each cue has an unequal contribution to the final solution (Davelaar, 2015; Wiles et al., 1991). Evidence supporting the WAM demonstrates that participants are more likely to select one of three cues as a primary cue to search in a typical convergent-thinking task (Smith et al., 2013). Massaro et al. (1991) also pointed out that "the only difference between the WAM and the FLMP is the nature of the integration algorithm – assumed to be compromising for the WAM and enhancing for the FLMP" (p. 280). However, the WAM does not specify the detail of the mechanism of the compromising process (Calvert et al., 2004), and the exact nonlinear integration process is still unclear (Davelaar, 2015).

A further complication in understanding of how cues are processed is the issue that boundaries between different models are not absolute (Davelaar, 2015). Most studies attempt to decide between a local or global search process (or serial vs. parallel search) without explicit investigation of the interaction between cues (Massaro et al.'s study is a notable exception). In general, multiple-cue interaction and serial versus parallel search are independent of each other. For example, although independent cue-processing is usually assumed in a standard parallel model (Houpt & Townsend, 2012); cues could interact with each other or may not be equally evaluated because the architecture and the relationship between cues are independent from each other (Altieri et al., 2017; Townsend & Nozawa, 1995). Eidels et al. (2011) outlined a framework for understanding how a parallel model can involve different interactions (i.e. facilitation or inhibition). Thus, it is necessary to explore the relationship between multiple-cue processing behavior.

Here, we apply a theory-driven methodology, Systems factorial technology (SFT; Townsend & Nozawa, 1995; Little et al., 2017), to measure the dual-cue relationship. This method allows investigation of whether the semantic search is serial or parallel along with an assessment of processing efficiency, thus clarifying the dual-cue interaction and understanding the possible mechanism of dual-cue processing efficiency in a convergent thinking search task. Recent work by Howard et al. (2020) focused on serial-parallel processes in multiply-constrained search (where the outcomes is determined by the intersection of multiple cues) and found some evidence for parallel processing but did not look at capacity. The present study focuses on processing efficiency and, to our knowledge, is the first application of capacity analysis to convergent thinking.

In this paper, we ask how dual cues change the efficiency of searching for a solution compared to single cues in a convergent thinking task. To assess this question, we examine a measure termed *workload capacity*, which, in general, denotes the variation of processing efficiency as workload changes.

In the present context, workload capacity refers to the manner in which processing efficiency changes as a function of the number of cues (Eidels et al., 2011; Houpt & Townsend, 2012; Little et al., 2017; Townsend & Nozawa, 1995). That is, as an additional cue is added, the rate of processing – or how quickly one can find an appropriate semantic term – may have a general increase or decrease (or no change). We analyzed the capacity of cue-combination under a conjunctive-rule condition (AND rule) that required participants to associate two different cues with a solution. During this process, a correct answer must be matched with both cues, which accords with the requirements of a multiply-constrained problem solving test. Going back to our opening example, this means that the correct answer (say, Mrs. Smith), must be matched with both cue 1 (Geography Teacher) and cue 2 (high school) before the system can provide a correct response.

Workload capacity can also inform researchers about whether there is an interaction between cues. Specifically for multiple-cue memory search, a supercapacity system could possess a positive interaction in which multiple cues facilitate mutual processing so that processing speeds up and efficiency increases. In a limited-capacity system, processing speed slows down as an additional cue is added. In the limited-capacity cases, cues may inhibit the evaluation of another source so that the processing efficiency reduces. Finally, in an unlimited-capacity system, speed and processing efficiency are not changed when an additional cue is added.

## Measuring Workload Capacity

According to Townsend and Wenger (2004), the capacity coefficient under the AND-rule condition is expressed as:

$$C_{AND}(t) = \frac{[K_L(t) + K_R(t)]}{K_{L,R}(t)},$$

where $K(t) = log[F(t)]$, $t > 0$. $F(t)$ is the cumulative distribution function (CDF), $F(t) = \mathrm{P}\{\text{Reaction time } (RT) \leq t\}$. $K_{L,R}(t)$ denotes the reverse cumulative hazard function (Chechile, 2011) of the redundant-cue trials, while $K_L(t)$ and $K_R(t)$ denote the same function for the two single-cue trials – one presented on the left and one presented on the right, respectively. Somewhat more intuitively, $C_{AND}(t)$ compares the response-times distribution when there are cues in both locations (denominator) to the expected maximum time distribution derived from each cue separately (numerator) under the assumption that processing is parallel and exhaustive and the assumption of context invariance (Otto & Mamassian, 2012). The derived maximum time distribution provides a baseline for comparing actual dual-cue performance. If $C_{AND}(t) > 1$, the processing is of supercapacity – dual cue response are faster than expected under parallel exhaustive processing; if $C_{AND}(t) < 1$, the processing is of limited capacity – dual cue responses are slower than expected under parallel exhaustive processing; if $C_{AND}(t) = 1$, capacity is unlimited – or as expected under the parallel exhaustive baseline. Throughout, $C_{AND}(t)$ was estimated using the SFT [R] package (Houpt et al., 2014).

$C_{AND}(t)$ is a function of time; hence, the parenthetical argument, t. To interpret $C_{AND}(t)$ across time, we additionally computed an upper and lower bounds on performance for unlimited-capacity independent parallel models (see Appendix A). Violation of either side of the bounds is strong evidence for rejection of unlimited capacity (upper bound for supercapacity and lower bound for limited capacity). Hence, the measurement of capacity along with these bounds provides a test for the level of interaction between cues. If the dual cues are processed in an integrative fashion for convergent thinking, we should expect supercapacity processing (that is, $C(t) > 1$, possible violations of the upper bound). Alternatively, if either limited or unlimited capacity processing is observed, considering parallel processing observed by Howard et al. (2020), the cues may operate in parallel with one of the cues predominating in the weights, or even operate thoroughly independently (see General Discussion for more details).

## Disjunctive OR-Rule Condition

We additionally assessed the workload capacity under a disjunctive, OR-rule condition to test the possibility of complete independence between cues. In the OR-rule condition, as soon as one of the cues has been processed and an answer has been generated based on that single cue, a decision can be made (a self-terminating stopping rule, Townsend & Nozawa, 1995). Therefore, we tested whether participants can process the two cues independently and in parallel when integration of cues is not necessary. To investigate capacity in this task, we computed the capacity coefficient for an OR rule:

$$C_{OR}(t) = \frac{H_{L,R}(t)}{[H_L(t) + H_R(t)]},$$

where $H(t) = -log[1 - F(t)]$, $t > 0$. $H_{L,R}(t)$ denotes the cumulative hazard function of the redundant-cue trials, while $H_L(t)$ and $H_R(t)$ denote the cumulative hazard function for two single-cue trials – Left and Right, respectively. The inferences for the capacity coefficient under the OR-rule condition is the same as those under the AND-rule condition (see Appendix A). Generally, we expect unlimited-capacity processing with cues independently processed in an OR-rule condition.

## Measuring Convergent Thinking

In addition to estimating the capacity measures, we also conducted a Chinese version of the Remote Associates Test (RAT; Mednick, 1962; 1968) to measure participants' ability of convergent thinking. The RAT is a common convergent thinking test (Nielsen et al., 2008), widely employed to measure a participant's ability to perform an effective logical search, exercise appropriate recognition, and make decisions that involve integration of divergent solutions into a response (Cropley, 2006; Jaarsveld & Lachmann, 2017). The RAT is also associated with creativity (Mednick, 1962) and insight (Schooler & Melcher, 1995). In a standard RAT, there are usually 30 test items, each of which contains three mutually unrelated cues (e.g., 'moon', 'dew', and 'comb'; Smith et al., 2013). Participants are required to solve the problem by searching for an answer that can fit all the given cues. Performance is assessed by measuring the number of correct answers. Here, we calculated correlation between RAT performance and the workload capacity of the dual-cue task. Workload capacity is often overlooked by studies of convergent thinking, especially the study of semantic processing (Parker, 2019), but is a basic property of all systems processing multiple elements and potentially plays an important role in the process of convergent thinking. The current study investigates workload capacity in a convergent thinking test, remote associations, and their relationship.

Altogether, our study aims to explore the possible mechanisms underlying convergent thinking. The present study investigates the dual-cue interaction process revealed by workload capacity. Second, our study examines the relationship between workload capacity and the performance of the convergent thinking test. Intuitively, people with supercapacity are expected to score higher on RAT, and there should be a positive relationship between workload capacity and RAT because supercapacity indicates high processing efficiency. Alternatively, it has been demonstrated that individuals with high RAT score are able to generate more candidate answers (Benedek & Neubauer, 2013; De Dreu et al., 2012), which may result in a consumption of processing resources. Thus, a potential negative correlation between workload capacity and RAT scores may be expected. Here, we tested the alternative hypotheses, and implications on the study of semantic search and convergent thinking process will be discussed.

# Experiment 1

In Experiment 1, participants performed a word production task by associating the given cue(s) with a related word. Similar to the RAT, in the conjunctive AND-rule condition of the word production task participants are required to combine two presented cues to search for a word that is related to both cues. Compared with the RAT, in which there is generally only one correct answer per test item, there may be more than one candidate answer. Thus, we required participants to answer using the first solution found which fit both cues. In addition, unlike the RAT, we did not force participants to generate a remote or creative solution.

In the word production task, we employed SFT analyses to assess the capacity coefficient, $C_{AND}$(t), for each participant. In addition, an operation span task (OSPAN; Turner & Engle, 1989) was adopted to measure participants' working memory capacity (WMC) in order to ensure that the individual differences are not a result of differences in working memory capacity (Benedek & Neubauer, 2013; De Dreu et al., 2012; Furley & Memmert, 2015; Hedblom, 2013). Participants also completed a Chinese version of the Remote Associates Test (CRAT; Huang & Chen, 2012) to measure their convergent thinking ability. The latter allowed the examination of the relationship between workload capacity and convergent thinking behavior.

## Method

### Participants

Forty-six participants (22 male and 24 female) participated in this experiment. Their ages ranged from 18 to 35 years old ($M$ =

22, $SD$ = 3.8). All were native Mandarin speakers born and raised in Taiwan and had normal or corrected-to-normal vision. Each received four points of introduction to psychology course or NTD 120 per hour after their participation. They signed a written informed consent prior to the experiment. The ethics approval for the study was obtained from the Ethics Committee of Department of Psychology at National Cheng Kung University, and the study was conducted in accordance with the approved guidelines and regulations.

### *Apparatus*

The experiment was programmed with E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). All the stimuli were displayed on a 19-in. CRT monitor (CTX VL951T) with a refresh rate of 85 Hz. The display resolution was 1024 (width) x 768 (height) pixels. The viewing distance was kept at 60 cm.

### *Stimuli, Design, and Procedure*

**Semantic Search Task.** The test cues were randomly selected from 78 adjectives that were divided into 6 categories (color, shape, texture, exterior, state, and taste). To ensure the representativeness of the test cues for each category, we asked an additional fifty participants[1] from the same subject pool, who did not participate in the formal experiment, to rate to what extent each test adjective belongs to the category using a 5-point Likert-Type scale. The rating results showed that the average score ranged between 4.0 to 4.9 (SDs < 1.1). Please see Appendix B for the word list and the rating results. In the formal experiment, the test cues were presented in Courier New font with a font size of 18 pt and were displayed in white on a black background.

There were 600 formal test trials, divided into 10 blocks (see Figure 1). One third of trials were redundant-cue trials, where two test cues were presented simultaneously. The remaining trials were single-cue trials, half of which contained a cue presented on the left-hand side of the screen (left cue) and half of which contained a cue presented on the right-hand side of the screen (right cue). We did not include no-cue trials. These often serve as essential catch trials in ordinary redundant-targets tasks but were not required in the current design since participants had to type in their complete response rather than select from alternatives. All three conditions were randomly intermixed within a block with equal presentation frequency. Before the formal experiment, 30 practice trials were conducted to ensure that the participants were familiar with the experimental instruction and procedure.

On each trial, a central fixation cross appeared for 500 ms, followed by a cue display presented until the participants indicated that they have an appropriate answer by pressing the "SPACE" bar to terminate the display. A text box then appeared, and the participants had to type their answer into it, later checked for correctness by the experimenters. In the redundant-cue trials, participants had to generate an answer that was related to both cues; in the single-cue trials, participants had to generate an answer related to that cue. RT was recorded from initial cue presentation to the spacebar press. Due to the limited number of cues, participants were required to generate different answers given the same cue(s); otherwise the repeated answer was considered incorrect. The intertrial interval was 1000 ms.

**Operation Span Task.** In the OSPAN task (see Figure 2), the participants were first required to decide whether an arithmetic equation was correct, followed by a to-be-remembered two-character Chinese word. Each trial had 2, 3, 4, 5, or 6 such arithmetic-and-storage presentations. As soon as the presentations finished, participants needed to retrieve the to-be-remembered words with a correct serial order immediately. Five presentation conditions (2/3/4/5/6) were equally assigned into a total of 15 trials all of which were randomly intermixed. Throughout the task each arithmetic equation or two-character word (i.e. each presentation) was presented once. Please see Appendix C for the list of Chinese words used in this task.

**Chinese Remote Associates Test (CRAT).** In the CRAT (Huang & Chen, 2012) there were 30 test items, each item composed of three cues and participants were required to generate an answer that can simultaneously relate to the three test cues. For example, if the cues were Taiwan (台灣), mainland (大陸), and barrier (屏障), the correct (and only) answer would be strait (海峽). The number of correct answers was recorded as the final score. The test has high reliability with Cronbach's alpha of .81, and the correlation between the CRAT score and scores for the insight problem test is .51, suggesting the CRAT has a good test validity of measuring one's convergent thinking ability (Huang & Chen, 2012; Huang et al., 2012).

## Results

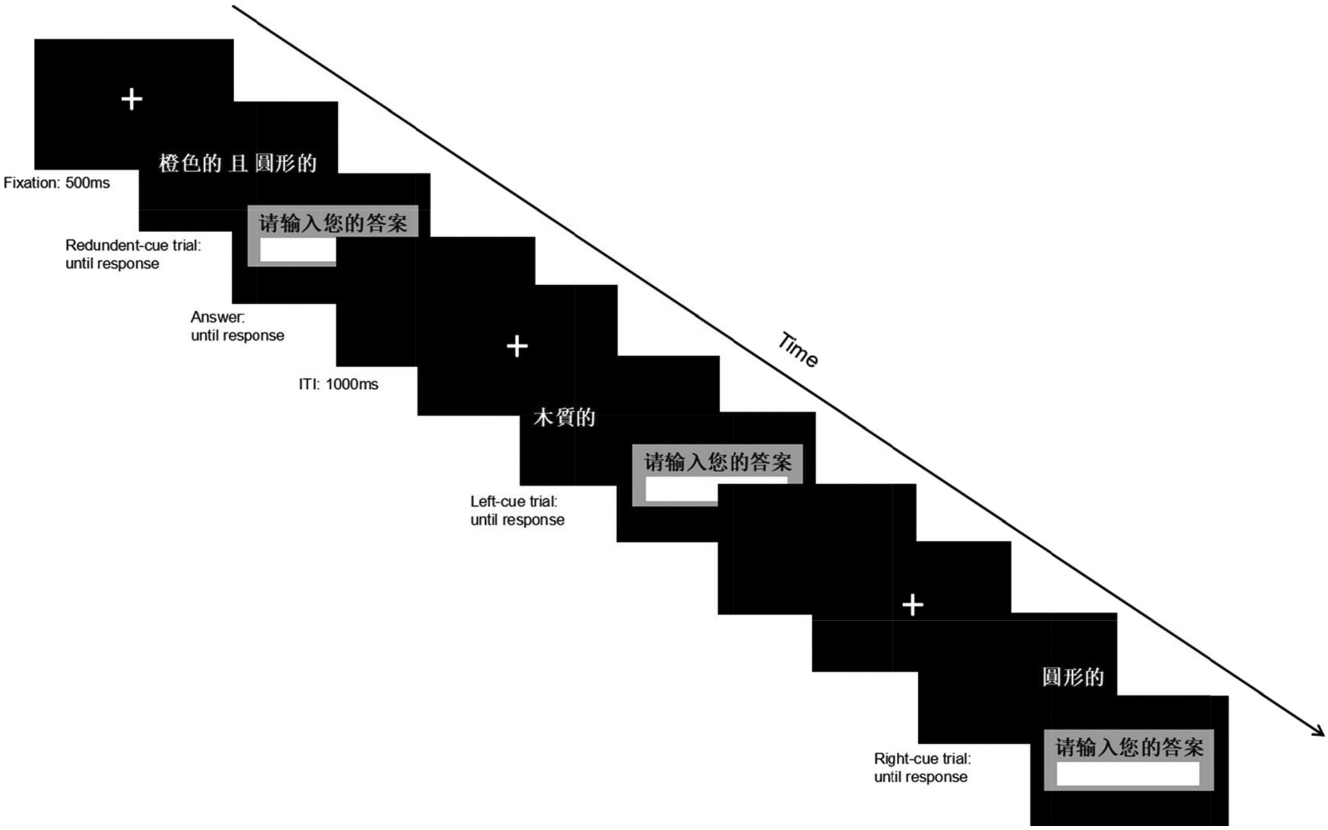
In the word production task, a response was considered correct when the answer was judged to be related to the cue(s) by the experimenter. For example, if the cue was 'circular', a correct answer would be the word related to circular (e.g., an 'orange') whereas a 'magic square' would be considered as an incorrect answer. If the participants responded "I have no idea", the trial would be also regarded as incorrect. In addition, exceptional ideas and subjective answers would be accepted, if the participants provided a reasonable explanation posthoc. For example, a participant responded to 'dwarf' with an unfamiliar name but subsequently explained this is the name of his cousin, who had been an undersized boy. This type of answer was accepted as correct, based on the idiosyncratic relationship.[2]

Throughout the experiment, because of a limited number of cues, we asked the participants to generate different answers to the same cue; if a participant repeated the same answer given the same cue(s), the response would also be excluded. Under this criterion, one participant was excluded due to low accuracy (65% in single-cue trial) with a large number of repeated answers or

[1] All the participants were the students at National Cheng Kung University who took Introduction to Psychology course. They completed the Likert-type scale questions to rate to what extent each test adjective belongs to the category and received one point of Introduction to Psychology course.

[2] We acknowledge that there are several objective methods to assess the correctness of an answer by analyzing the associative strength between the generated answer and the given cues using the associative norm (e.g., De Deyne et al., 2013). However, there is no appropriate norm to analyze the associative strength between the generated answer and redundant-cues especially when the two cues are unrelated with each other. In addition, instead of semantic relevance and similarity, we also consider the idiosyncratic relationship into account while assessing the correctness.

**Figure 1**
*An Illustration of Three Experimental Trials: The Redundant-Cue, Left-Cue Trial, and Right-Cue Trial*




*Note.* In this example, the redundant cue is "orange" and "circular" (橙色的 且 圓形的), the left cue is "woody" (木質的), and the right cue is "circular" (圓形的).

irrelevant answers with no reasonable explanations. For example, she associated "rhombic and fat" with "Alice in Wonderland" and explained that there was no reason and it was just a feeling. Moreover, she repeated this answer in different single-cue and redundant-cue trials. Thus, we considered her data invalid.

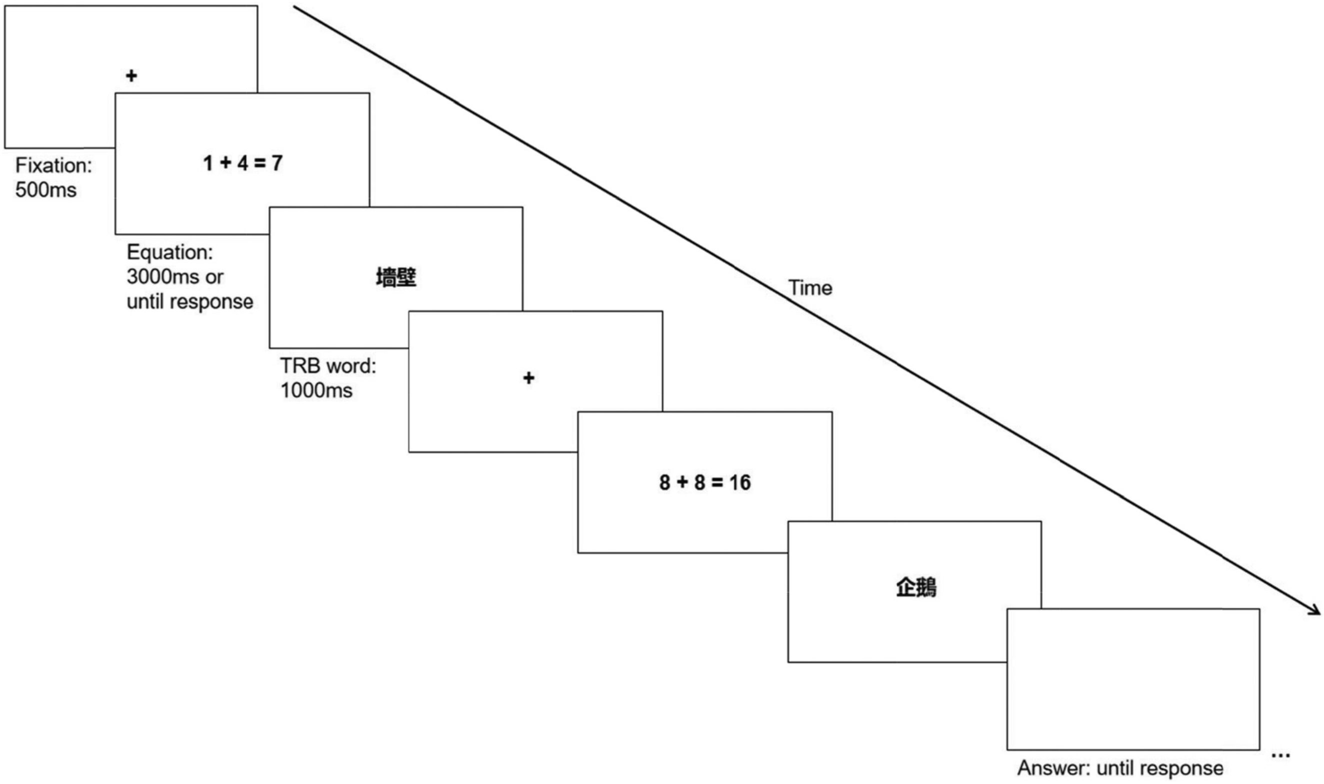
Based on all the correct responses, mean RT, and standard deviation of each condition for each participant were calculated. Participants with mean RTs beyond 3 standard deviations away from the group mean were excluded as outliers. In addition, participants with WMC or CRAT score above or below 3 standard deviations from the group mean were excluded. Under this criterion, one participant's mean RTs for both the single-cue and redundant-cue trials were above 3 standard deviations, so his data were excluded from further analysis.

Across all participants, a total of 4.88% data points were excluded. Table 1 presents mean performance for the remaining data in Experiment 1. Overall, single-cue trials were significantly less accurate than the redundant-cue trials, $t(43) = -6.06$, $p < .001$, $d = 1.27$ and significantly faster, $t(43) = -10.68$, $p < .001$, $d = .50$. This relationship was not caused by a speed–accuracy tradeoff but by the exclusion of the repeated answers, due to which a large number of single-cue trials (6% of single-cue trials) were deleted even though they were correct. If these answers remained, the mean processing accuracy in both the single-cue and redundant-cue trials would have been higher, .996 and .987, respectively.[3]

Figure 3A plotted the $C_{AND}(t)$ for all the participants in Experiment 1, showing a similar trend. The capacity coefficient function for each participant fluctuated between the capacity bounds for medium RTs (see Appendix A), but above the upper bound for the faster RTs. It also violated the lower bound for the slower RTs. To explore whether individual differences exist in processing efficiency despite a general decreasing trend, we adopted a z-transformation to linearly transform each individual function to a capacity z-score (CZ) (Houpt & Townsend, 2012). Here, CZ and C(t) capture different properties of the workload capacity. While C(t) reflects general trends and cognitive dynamics across all RTs, CZ shows the mean result and average performance at these time points. The CZ results showed individual differences in the workload capacity (i.e. processing efficiency): about half of the participants (21 of 44) exhibited nonsignificant CZ, suggesting unlimited capacity. Twelve participants' CZ was significantly less than 0, suggesting limited capacity, while for

[3] Here, single-cue trials were significantly more accurate than the redundant-cue trials, $t(43) = 2.95$, $p < .001$, $d = 0.57$ and significantly faster, $t(43) = -10.80$, $p < .001$, $d = 1.51$, because it was more difficult to find out a correct answer based on dual cues than based on one given cue.

**Figure 2**
*An Illustration of the Experimental Procedure of the OSPAN Task (Take 2 Presentations for Example)*




*Note.* The to-be-remembered words (墙壁, 企鹅) denote wall and penguin.

eleven participants CZ was significantly greater than 0, suggesting supercapacity. At the group-level analysis CZ was consistent with unlimited capacity.

We further adopted functional principal component analysis with varimax rotation (fPCA; Burns et al., 2013) to identify the components that resulted in the capacity differences. These component functions can be used to display increase or decrease trends compared with the mean capacity function, and each component can explain that percentage of variation. As shown in the left panel of Figure 4A, the first two principal components (PCs) accounted for 77% of variance (PC 1: 65% and PC 2: 12%). PC 1 displayed an overall increase in workload capacity compared to the mean capacity while PC 2 represents the slope of the capacity function. Positive PC 2 scores reflect more positive slope than average and negative scores reflect more negative slope.

In order to examine the relationship between workload capacity and convergent thinking behavior, we used two approaches. The first, the extreme-group approach, has been widely used to analyze continuous variables (Preacher et al., 2005). We employed it to emphasize the differences between high- and low-CRAT groups on the basis of top and bottom 30% quantile. The second approach was to apply linear regression, which revealed clear relationship between the capacity and the convergent thinking ability.

In the extreme group approach, participants were split into high-CRAT ($M$ = 21.87, $SD$ = 1.64, $N$ = 15) and low-CRAT groups ($M$ = 14.62, $SD$ = 1.66, $N$ = 13) on the basis of top and bottom 30% quantile. The difference in CRAT scores across groups was significant, $t(25.35) = 11.57$, $p < .001$, $d = 4.39$. An independent-samples t-test revealed a significant difference in the CZ across the high- and low-CRAT groups, $t(23.95)$ =

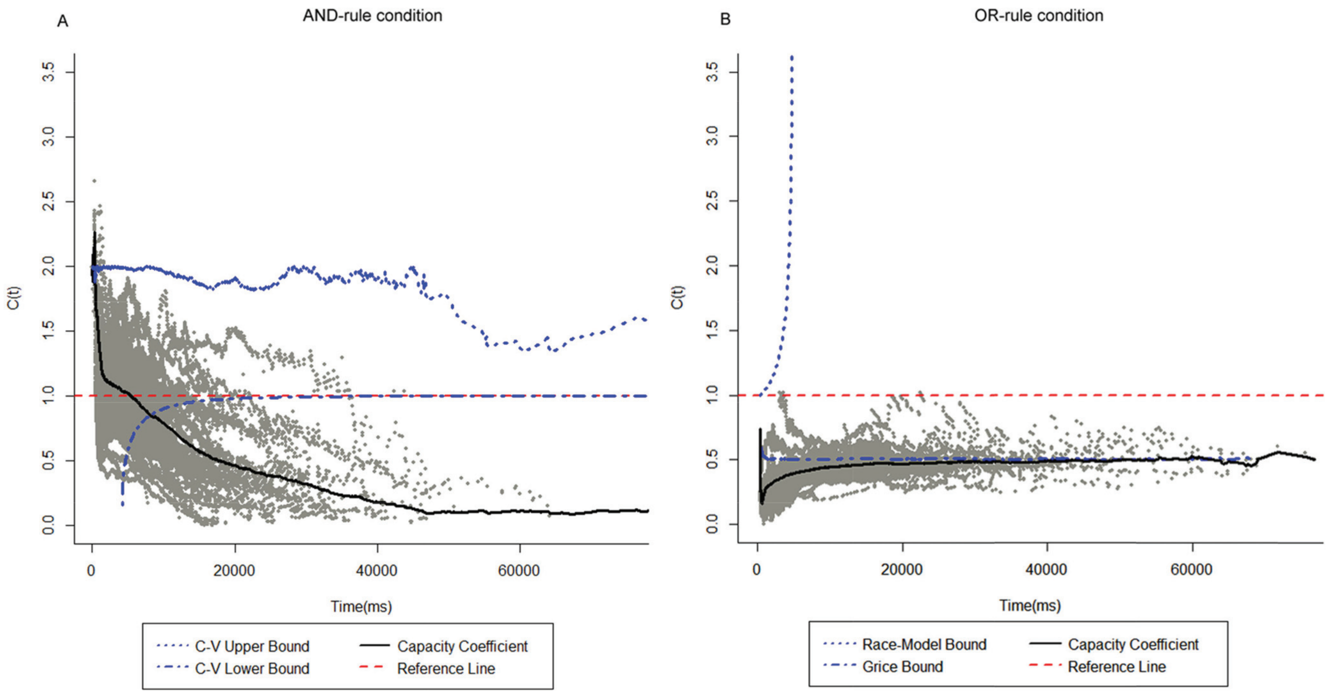
**Table 1**
*M Accuracy (ACC) and RT (in Sec) of the Redundant-Cue and Single-Cue Trials, CRAT, and WMC Scores in Experiment 1*

| | Redundant-cue | | Single-cue | | Left-cue | | Right-cue | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ = 44 | ACC | RT | ACC | RT | ACC | RT | ACC | RT | CRAT | WMC |
| $M$ | 0.98 | 11.63 | 0.93 | 6.09 | 0.93 | 6.16 | 0.94 | 6.01 | 18.11 | 31.25 |
| $SD$ | 0.02 | 4.73 | 0.05 | 2.41 | 0.05 | 2.45 | 0.05 | 2.39 | 3.25 | 10.73 |

*Note.* CRAT = Chinese Remote Associates Test; WMC = working memory capacity.

**Figure 3**

*Plot of the C(t). A, $C_{AND}(t)$ With C-V Bounds in Experiment 1. B, $C_{OR}(t)$ With Race-Model Bound and Grice Bound in Experiment 2. For Both A and B, the Black Line is the Group Capacity While the Dark Gray Dots are the C(t) for Each Individual*




*Note.* See the online article for the color version of this figure.

$-3.96$, $p < .001$, $d = 1.48$ but no significant difference in their WMC scores, $t(25.45) = .09$, $p = .93$, $d = .03$, suggesting high-CRAT participants had lower workload capacity than the low-CRAT participants. Figure 5A compares the distributions of CZ from the two groups.

The above findings demonstrated there were differences in the overall capacity level across the CRAT groups – with higher capacity scores for the low CRAT group. These results imply a negative relationship between workload capacity and convergent thinking ability. Pearson correlation coefficient confirmed a significant negative correlation between capacity and CRAT[4], $r(42) = -.44$, $N = 44$, $p = .003$. The negative correlation was also observed between the loadings of PC 1 and CRAT scores, $r(42) = -.39$, $N = 44$, $p = .009$, but did not reach the significance level between the loadings of PC 2 and CRAT scores, $r(42) = .16$, $N = 44$, $p = .31$. No significant relationship was observed between WMC scores and CRAT scores, $r(42) = .14$, $N = 44$, $p = .36$, nor between CZ and WMC scores, $r(42) = .03$, $N = 44$, $p = .83$. In order to remove the effect of WMC, a partial correlation analysis was used to analyze correlation between CZ and CRAT scores. The negative correlation still held, $r(42) = -.45$, $N = 44$, $p = .003$. Importantly, there was no significant difference between the correlation and the partial correlation coefficients ($p = .52$). Also, no significant correlation was observed between RT on single-/dual-cue trials and CRAT ($r = -.07$, $p = .26$ and $r = .09$, $p = .55$, respectively). The negative correlation suggests a simple linear model predicted CRAT based on capacity. Figure 6A summarizes the correlation results.

## Discussion

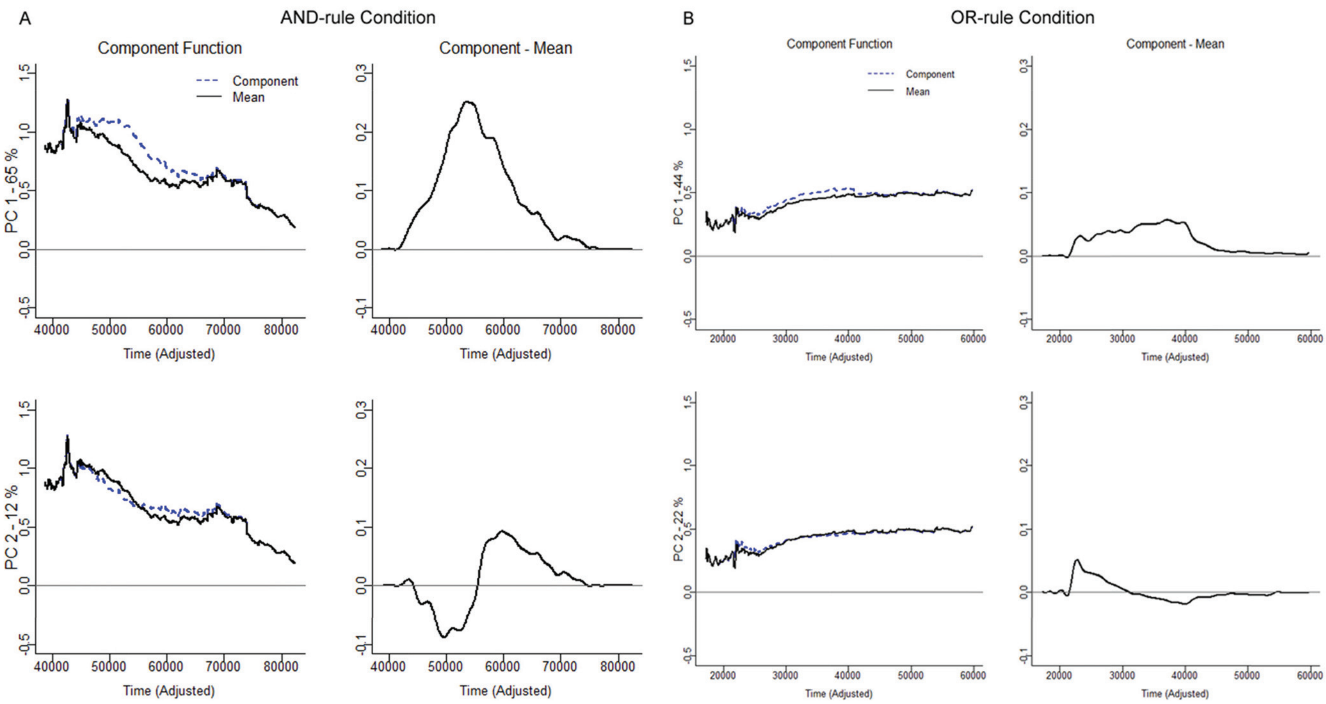
In general, all empirical $C_{AND}$(t) plots presented a steadily monotonically decreasing trend. Although all observed $C_{AND}$(t) functions displayed a very similar trend, the CZ findings revealed large individual differences in processing efficiency, which also suggests the multiple-cue combination in convergent thinking is not absolutely limited to one fixed pattern (Davelaar, 2015). The extreme group approach further demonstrated there were significant differences in workload capacity scores (i.e. processing efficiency) between high- and low-CRAT groups. Specifically, about half of the participants exhibited unlimited capacity, suggesting adding an additional cue did not change the processing efficiency. That is, dual cue processing might not interact with each other. For the other eleven participants with supercapacity, the processing of a single cue speeds up as an additional cue was provided, consistent with the prediction of the enhancing integration model. That is, the dual cues can facilitate each other with a positive between-channel interaction during the information accumulation process (Wenger & Townsend, 2001).

By contrast, limited-capacity processing, observed for the remaining twelve participants, could be the consequence of either of three likely mechanism. First, a standard serial processing with limited

[4] Even when we included the repeated answers into the analysis, the results remained consistent. The negative correlation between CRAT and CZ still existed, $r(42) = -0.45$, $N = 44$, $p = .002$, and no significant difference was observed between the correlation coefficient with and without repeated answers ($p = .52$).

**Figure 4**

*The fPCA Results. A, The fPCA Results Under AND-Rule Condition. B, The fPCA Results Under OR-Rule Condition. For Both A and B, the Left Panel Shows the Principal Component Function (Blue Dashed Line) and Mean Capacity Function (Black Line); the Right Panel Shows the Contrast Functions Calculated From the Mean Capacity Function Subtracted From the First/Second Principal Component Function*




*Note.* See the online article for the color version of this figure.

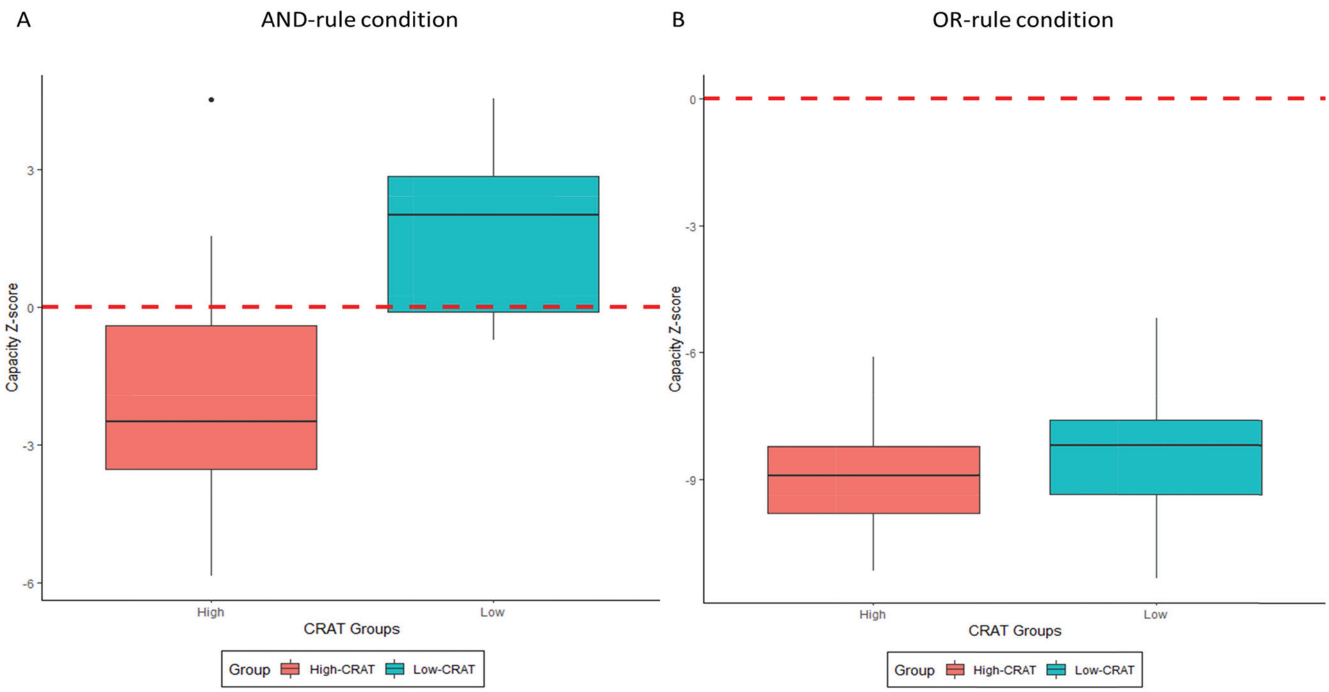
capacity (Townsend & Ashby, 1983) suggests dual cues can be processed sequentially and serially. That is, participants with limited capacity may process one cue at a time. Second, a parallel model with inhibitory interaction (Eidels et al., 2011) suggests dual cues can be processed simultaneously and in parallel, but inhibit each other. Participants may process both cues together, but these cues interact in a negative way due to competition for limited processing resources or semantic interference until one is assigned to a dominant weight. Third, there might exist a change in the rate of processing between the redundant-cue trials and the single-cue trials, which is regarded as a violation of context invariance assumption (Otto & Mamassian, 2012; Yang et al., 2018; Yang et al., 2019). Context invariance holds, for example, when the left cue is processed at the same rate regardless of whether it is presented alone or in the context of the redundant-cue trial. If the left cue is processed faster when presented alone and processed slower when presented with the right cue, then the dual-cue processing may still be independent and parallel but will show limited capacity because of the change in processing rate.

A possible cause of different capacity types with an overall similar trend is the availability of the processing resources. According to adaptive control of thought model (ACT model; Anderson, 1983); cues in LTM (Flexser & Tulving, 1978; Massaro et al., 1991) automatically activate some neighboring or related semantic nodes, as in the spreading-activation model (Collins & Loftus, 1975), and some of the nodes above the critical threshold are temporarily loaded into working memory as candidate answers (Cantor & Engle, 1993). When the direct effect of WMC is excluded, the number of candidate answers loaded becomes important for workload capacity. Generally, for each individual, WMC (and hence the amount of available processing resources) is fixed. Therefore, the more loaded candidates are with information, the less available resources they have for processing; this leads to the displayed lower processing efficiency. According to previous studies (Benedek & Neubauer, 2013; De Dreu et al., 2012), creative people can produce more candidate answers based on a given cue, suggesting that they can load more candidates into working memory for successful retrieval. This may deplete available processing resources, resulting in limited capacity. On the other hand, unlimited and supercapacity responses suggest some people have enough processing resources to process dual cues due to less candidate answers loaded.

The number of candidate answers loaded also confirms our prediction of the negative relationship between workload capacity and the CRAT. Individuals who are able to produce more related candidates according to the given cue have higher CRAT scores (Benedek & Neubauer, 2013; De Dreu et al., 2012). That is, individuals who obtain high scores on CRAT have more related candidates activated above the critical threshold. This may exceed the WMC, causing a decrease in available processing resources and thereby presenting low workload capacity. Conversely, individuals who score low on the CRAT show high workload capacity because, potentially, there are fewer candidates in their working memory leaving sufficient resources to other processes, manifesting high workload capacity. This implies that individuals who perform well in CRAT can activate more related candidates over the critical threshold than those who perform more poorly in CRAT,

**Figure 5**

*Boxplot of CZ Values for the High- and Low-CRAT Groups. A, Boxplot of CZ Values for the High- and Low-CRAT Groups Under AND-Rule Condition. B, Boxplot of CZ Values for the High- and Low-CRAT Groups Under OR-Rule Condition. For Both A and B, the Red Dashed Line at 0 is the Reference Line of Unlimited Capacity*




*Note.* See the online article for the color version of this figure.

in accord with previous studies (Benedek & Neubauer, 2013; De Dreu et al., 2012; Dietrich, 2004; Song et al., 2011). The fPCA findings suggested that the most significant contributor to enhanced workload capacity was negatively correlated with CRAT, conforming an inverse relationship between workload capacity and convergent thinking.

The results of Experiment 1 suggest multiple cues may be combined via a parallel interaction model. Yet, unlimited capacity could also be the consequence of independent parallel processing with recruitment of resources on redundant-cue trials (Altieri et al., 2017; Townsend & Nozawa, 1995). That is, dual cues may follow either completely independent processing (no interaction) or an interaction-parallel model in which the cues interact with each other to some degree or another, but the interaction between cues does not influence the cue-processing efficiency. In order to investigate whether dual-cues are operating independently or not, we adopted in Experiment 2 the OR-rule task to force the participants to search for an answer based on only one of the cues. If the interaction between the dual cues is not necessary for convergent thinking, participants should be able to generate an answer based on either cue alone, and a negative relationship between workload capacity and convergent thinking should hold, as observed in Experiment 1.

## Experiment 2

We used the same procedure as in Experiment 1 with a fresh group of participants but with an OR task, where only one cue was needed to produce a correct response, even if two cues were present. We calculated capacity and z-score capacity (CZ) and correlated the CZ with CRAT scores as we did in Experiment 1. If the interaction between the dual cues is not necessary for the process of convergent thinking, participants should be able to generate an answer in terms of either cue. In this case, a negative relationship between workload capacity and CRAT performance should be observed in an OR-rule condition as in the AND-rule condition in Experiment 1.

### Method

#### *Participants*

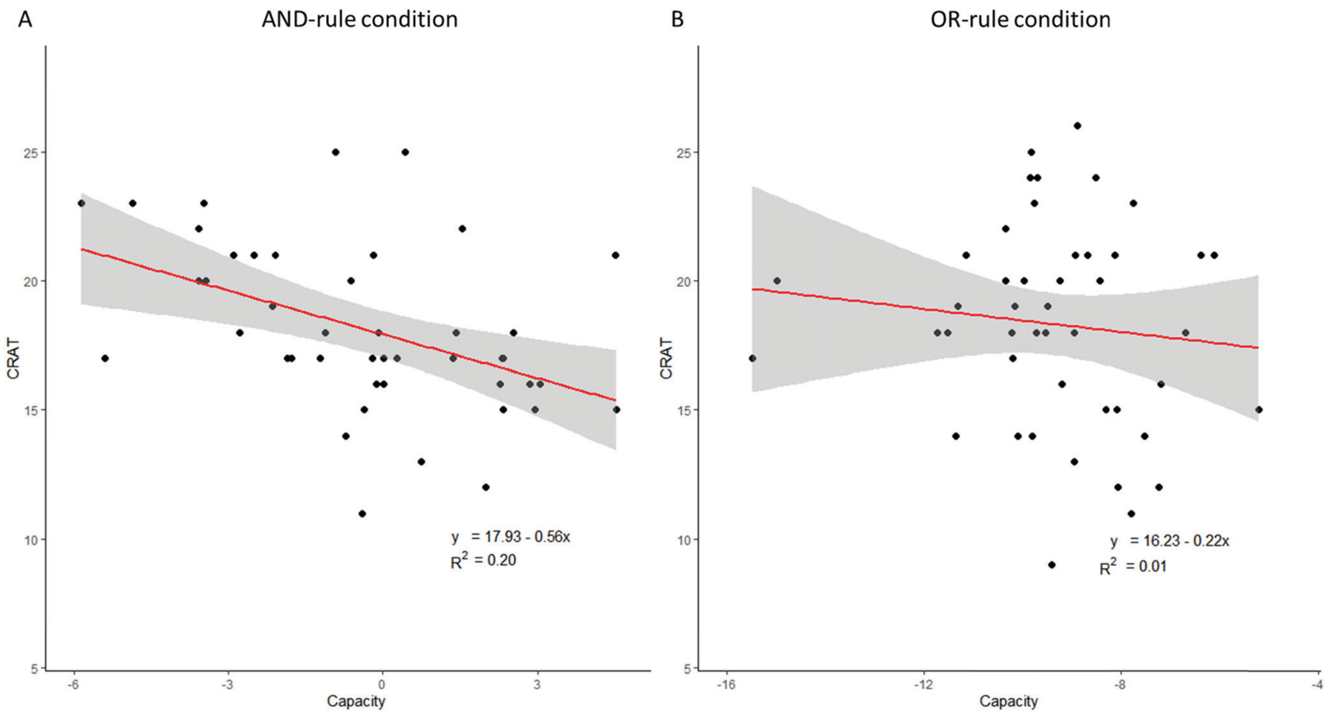
Forty-five native Mandarin participants in Taiwan (24 males and 21 males) whose age ranged from 18 to 35 years, with a mean age of 23 and a standard deviation of 3.3. All had normal or corrected-to-normal vision. The participants received 4 points of introduction to psychology or NTD 120 per hour after the experiment.

#### *Stimuli, Design, and Procedure*

The design of Experiment 2 was the same as Experiment 1 except for the instructions on redundant-cue trials. In the redundant-cue trial of Experiment 2, participants were required to respond with an associate of either of the cues that were presented. To illustrate, in Experiment 1, when participants saw ‘circular’ and ‘sweet’, the answer had to match both ‘circular’ and ‘sweet’ (e.g., ‘orange’). In Experiment 2 when participants saw ‘circular’ or ‘sweet’, a correct answer could be associated with either side alone (e.g., ‘sun’, or ‘honey’). All of the participants also completed the CRAT.

**Figure 6**

*Scatter Plot of Capacity Z-Scores (CZ) Versus CRAT Scores. A, Scatter Plot Showing the Relationship Between CZ and CRAT Under AND-Rule Condition. B, Scatter Plot of CZ Versus CRAT Under OR-Rule Condition. For Both A and B, the Red Line is the Regression Line, and the Gray Area Shadows the 95% Confidence Interval*




*Note.* See the online article for the color version of this figure.



### Data Analysis

Data in Experiment 2 were filtered using the same criteria as in Experiment 1. However, because of OR-rule, responses on redundant-cure trials were considered correct if they matched either of the cues.

## Results

No participants were excluded in Experiment 2. Figure 3B summarizes the capacity coefficient $C_{OR}(t)$ functions for all the participants. The $C_{OR}(t)$ for the group showed the function fluctuated below the reference line and the Grice bound for all t. This supports an inference of severely limited capacity. In addition, analysis of the z-transformed capacity coefficient function showed all participants displayed limited capacity, suggesting that an added cue can interfere with processing efficiency. Figure 4B shows results of the fPCA. The left panel shows PC 1 and PC 2 accounted for 44% of and 22% of variance, respectively. PC 1 displayed a slight increase in workload capacity around 20,000 ms, and PC 2 indicates a shift in the relative capacity values across time, that is, a change in the slope as with PC 2 in Experiment 1. Table 2 presents the descriptive statistics from Experiment 2.

We then conducted the same correlation and extreme group analyses used in Experiment 1. The correlation between CZ and CRAT scores was not significant (see Figure 6B), $r(43) = -.11$, $N = 45$, $p = .47$ even when WMC was controlled for, $r(43) = -.11$, $N = 45$, $p = .47$. Also, there were neither significant correlations between the loadings of PC 1 and CRAT scores, $r(43) = -.15$, $N = 45$, $p = .32$ nor between the loadings of PC 2 and CRAT scores, $r(43) = .27$, $N = 45$, $p = .08$. We split the participants into high-CRAT ($M = 22.64$, $SD = 1.74$, $N = 14$) and low-CRAT groups ($M = 13.57$, $SD = 1.99$, $N = 14$) as in Experiment 1, and the difference in CRAT scores across groups was significant, $t(25.54) = 12.85$, $p < .001$, $d = 4.86$. However, no significant difference was observed in the CZ across the high- and low-CRAT groups (see Figure 5B), $t(25.90) = -.74$, $p = .47$, $d = .28$.

## Discussion

In Experiment 2, we mostly observed limited capacity. This result is consistent with some published studies using SFT in an OR task to probe the processing of multiple signals. For example, Innes-Ker and Townsend (2003) observed supercapacity only under the conjunctive AND-rule condition but limited capacity under the disjunctive OR-rule condition. To account for the limited-capacity findings, one plausible explanation is that in the OR-rule condition participants process redundant cues in a serial fashion and choose one to process it first. This selection process is slow compared to just processing one cue so ends up being limited capacity. When both cues are needed under the AND-rule condition, there is no need for cue selection and this results in supercapacity or unlimited capacity. In fact, Townsend and Nozawa (1995) seminal OR application of the capacity coefficient and its

**Table 2**
*M Accuracy (ACC) and RT (in Sec) of the Redundant-Cue and Single-Cue Trials, CRAT, and WMC Scores in Experiment 2*

| | Redundant-cue | | Single-cue | | Left-cue | | Right-cue | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* = 45 | ACC | RT | ACC | RT | ACC | RT | ACC | RT | CRAT | WMC |
| *M* | 0.93 | 8.03 | 0.92 | 7.15 | 0.93 | 7.20 | 0.93 | 7.11 | 18.31 | 33.13 |
| *SD* | 0.05 | 3.75 | 0.05 | 3.57 | 0.05 | 3.60 | 0.05 | 3.38 | 3.95 | 10.56 |

*Note*. CRAT = Chinese Remote Associates Test; WMC = working memory capacity.

more recent replication by Eidels et al. (2015) had both reported unlimited to moderately limited capacity, above the Grice bound and above the capacity estimates reported in our Experiment 2. Notably, although these studies had employed two signals from the same domain, presumably competing for limited resources, the stimuli were plain perceptual stimuli (bright dots on dark background), lacking semantic processing and thus not 'costing' the cognitive system enough to exhaust its resources.

The present results indicate that a combination of cues is necessary in the process of convergent thinking, suggesting that cues cannot be independently processed throughout. They might operate independently before access into memory but must be combined to allow accurate retrieval of an answer. This is consistent with previous findings showing that people performed better by combining two cues than by just using a single cue (Massaro et al., 1991; Rubin & Wallace, 1989). It further suggests that the AND-rule condition, but not OR-rule, can capture the processes underlying convergent thinking during which multiple cues should interact with each other rather than operate with complete independence.

## General Discussion

The current study informed our understanding of the cognitive processes underlying multiply-constrained problem solving and convergent thinking. First, the overall trend of C(t) functions reveals a dynamic cue-interaction pattern during the search process. Second, individual differences in workload capacity are related to the final result of the convergent thinking. For the first issue, the overall trend with different capacities might not support a single processing model. In that sense, it is different from previous studies which argue that cues operated in terms of one fixed pattern (e.g., Massaro et al., 1991; Rubin & Wallace, 1989; Smith et al., 2013). In the study of Smith et al. (2013), the cues followed a weighted averaging model indicating that one cue is assigned with a dominant weight. In the study of Massaro et al. (1991), however, cues operate by enhancing integration that involves facilitatory interaction.

We believe the inconsistency between the current results and those observed in Smith et al. (2013) could arise from the use of different dimensions or quantities of cues. Smith et al. adopted a standard RAT in which all three cues were compared on their semantic similarities (or only on a semantic dimension). It is possible that resource consumption occurs under the same dimension, which certainly arouses fierce competition as cues are added. If this is the case, it would be reasonable to observe a difference between a dual-cue convergent task and the standard RAT. On the other hand, Massaro et al. (1991) and Rubin and Wallace (1989) adopted dual-cue tasks in which dual cues were from different dimensions (e.g., semantic dimension: 'a mythical being' and rhythmical dimension: 'rhymes with ost'). This suggests resource consumption occurs under different dimensions. Multiple cues from different dimensions provide each other with complementary features and are processed in a facilitatory-interactive manner. Generally, processing efficiency varies with different dimensions or quantities of the cues, which determines how multiple cues should be processed.

Differences between memory and convergent thinking studies may also cause the inconsistencies observed. Smith et al. (2013) and our present study both focus on multiply-constrained problem solving, whereas most previous studies focused on memory retrieval (Flexser & Tulving, 1978; Jones, 1976; Massaro et al., 1991; Wenger, 1999). Although retrieval is inevitably involved in the process of multiply-constrained problem solving, the biggest difference between them is joint memory (Smith et al., 2013). For some memory retrieval, especially for episodic memory, some cues are considered as contextual cues. They are usually learned together and retrieved together. However, for multiply-constrained problem solving, simultaneously acquiring all the cues is a small probability event, suggesting there is rarely a preexisting joint memory representation. Thus, the overlap of these cues is significant. The cues in multiply-constrained problem solving task are less likely to follow a completely independently parallel strategy, though they may operate independently before integration.

The second issue we addressed is that individual differences in workload capacity can account for the individual differences in convergent thinking. Specifically, cues access LTM and automatically activate related feature vectors (Anderson, 1983; Cantor & Engle, 1993; Collins & Loftus, 1975). At this time, workload capacity may work to decide which candidates can reach the threshold into working memory for keeping the activation to form the search patch (Cantor & Engle, 1993). Most previous studies assume that creative people should be able to load more candidates in working memory compared with uncreative people, which makes them generate more candidate responses (Benedek & Neubauer, 2013; Coney & Serna, 1995; De Dreu et al., 2012; Mednick et al., 1964). This assumption implies a positive relationship between WMC and creativity, but was not supported by existing evidence. So far few studies have observed any direct relationship between WMC and creativity (Benedek & Neubauer, 2013; De Dreu et al., 2012; Furley & Memmert, 2015; Hedblom, 2013). We also failed to obtain a significant correlation between them.

Here exists another possible explanation that the comparison may occur within-subject, rather than between the two different creative groups. Specifically, creativity is mainly characterized by high associative fluency, suggesting more related answers within a limited time according to a given cue (Benedek & Neubauer, 2013; Mednick, 1962). Also, creative people are more inclined to provide uncommon and rare answers compared with uncreative people (Benedek & Neubauer, 2013; Mednick, 1962). On the basis of the two characteristics we speculate that, for uncreative people, the search patches may be produced in a limited local range that requires a smaller number of resources, thereby displaying high workload capacity. Consequently, their answers are fairly high frequency (common) and neighbors around the given cue. From these patches, uncreative people decide on one of those candidates as a fitting solution. However, the retrieved candidate is not necessarily correct. The worst outcome is that they still fail to find out the correct answer after local patch depletion. This suggests they have to turn to global exploration, repeating the previous steps in order to form new search patches (Hills et al., 2015). This may lead to fewer correct answers within fixed time limits.

In contrast, the search patches for creative people may scatter in a wide range, which may burden working memory capacity, thereby displaying low workload capacity. However, wide-range patches suggest that the probability of involving the correct answer is increasing. This is why creative people are more likely to find out the correct answer within a fixed time limit.[5] A possibility exists that, for some people with higher creativity, the correct answer is decided as soon as the patches are formed, implying parallel search. Furthermore, if one of these patches obtains the dominant weight during the search-patch formation, parallel processing with limited capacity would be observed. Note that architecture and capacity are independent from each other (Altieri et al., 2017). However, by assessing the workload capacity, it is possible to infer a system's processing dependency or even to make an inference about the architecture when context invariance assumption holds. For example, independent cue-processing usually involves a standard parallel model (Houpt & Townsend, 2012). This is not to say, we repeat, that the cues in parallel model must be equally evaluated or must not interact each other because the architecture and the relationship between cues are independent from each other. Yet, the present study did not directly measure context invariance. Generally, search patches are considered to be formed in parallel while search process is sequential from one patch to another, or item by item in a local patch (Hills et al., 2015). This provides an elegant explanation, in our view, for the unlimited capacity observed in our study.

Overall, when the effect of working memory capacity is excluded, workload capacity becomes the key point to account for the convergent thinking performance, which proposes another possible explanation. In this explanation, however, we cannot directly compare how many candidate solutions have been loaded into working memory between the high workload capacity Group and the low workload capacity group. The cues, when processed by the high workload capacity group, may activate more candidates than the same cues processed by the low workload capacity group. On the other hand, we have only measured capacity and cannot make sure inferences about dependency and architecture. Nonetheless, we do demonstrate the negative relationship between the cue-processing efficiency and the performance on convergent thinking, and provide a blueprint for future study.

## Conclusion

We assessed workload capacity in a word production task and found support for individual differences in the model underlying the processing and integration of dual cues. In addition, we found performance on convergent thinking becomes worse as workload capacity improved. The results highlighted the role of workload capacity in convergent thinking task and shed light on findings of previous studies, which provide a new perspective for the mechanisms underlying convergent thinking.

[5] Here, we excluded another explanation that people with low efficiency to be faster to switch from one candidate to another so that they can perform more flexible and better on creativity because no correlation was observed between CRAT and RT in the single-cue trials, $r(42) = -0.17$, $N = 44$, $p = .26$, nor in redundant-cue trials, $r(42) = 0.09$, $N = 44$, $p = .56$. Even when the repeated answers retained, the correlations were still not significant, $r(42) = -0.18$, $N = 44$, $p = .24$ for single-cue trials and $r(42) = 0.09$, $N = 44$, $p = .55$ for redundant-cue trials, respectively. A previous study (Hommel, 2015) also claimed that the persistence-heavy control state, rather than a flexibility-biased state, contributed to convergent thinking task. Hence, we believe that the key factor is the number of candidate answers loaded in working memory.

## References

Altieri, N., Fific, M., Little, D. R., & Yang, C.-T. (Eds.). (2017). Historical foundations and a tutorial introduction to systems factorial technology. *Systems factorial technology: A theory-driven methodology for the identification of perceptual and cognitive mechanisms* (pp. 3–25). Elsevier Science Publishing Co Inc. https://doi.org/10.1016/B978-0-12-804315-8.00002-1

Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.

Anderson, N. H. (1974). Algebraic models in perception. In E. Carterette & M. Friedman (Eds.), *Handbook of perception II* (pp. 215–298). Academic Press.

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

Burns, D. M., Houpt, J. W., Townsend, J. T., & Endres, M. J. (2013). Functional principal components analysis of workload capacity functions. *Behavior Research Methods*, *45*(4), 1048–1057. https://doi.org/10.3758/s13428-013-0333-2

Calvert, G. A., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*. MIT Press.

Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1101–1114. https://doi.org/10.1037/0278-7393.19.5.1101

Chechile, R. A. (2011). Properties of reverse hazard functions. *Journal of Mathematical Psychology*, *55*(3), 203–222. https://doi.org/10.1016/j.jmp.2011.03.001

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology*, *38*(1), 35–58. https://doi.org/10.1006/jmps.1994.1002

Coney, J., & Serna, P. (1995). Creative thinking from an information processing perspective: A new approach to Mednick's theory of associative hierarchies. *The Journal of Creative Behavior*, *29*(2), 109–132. https://doi.org/10.1002/j.2162-6057.1995.tb00740.x

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

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

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, *45*(2), 480–498. https://doi.org/10.3758/s13428-012-0260-7

De Dreu, C. K., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin*, *38*(5), 656–669. https://doi.org/10.1177/0146167211435795

Vries, H. B., & Lubart, T. I. (2019). Scientific Creativity: Divergent and Convergent Thinking and the Impact of Culture. *The Journal of Creative Behavior*, *53*(2), 145–155. https://doi.org/10.1002/jocb.184

Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic Bulletin & Review*, *11*(6), 1011–1026. https://doi.org/10.3758/BF03196731

Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology*, *55*(2), 176–190. https://doi.org/10.1016/j.jmp.2010.11.003

Eidels, A., Townsend, J. T., Hughes, H. C., & Perry, L. A. (2015). Evaluating perceptual integration: Uniting response-time-and accuracy-based methodologies. *Attention, Perception & Psychophysics*, *77*(2), 659–680. https://doi.org/10.3758/s13414-014-0788-y

Flexser, A. J., & Tulving, E. (1978). Retrieval independence in recognition and recall. *Psychological Review*, *85*(3), 153–171. https://doi.org/10.1037/0033-295X.85.3.153

Furley, P., & Memmert, D. (2015). Creativity and working memory capacity in sports: Working memory capacity is not a limiting factor in creative decision making amongst skilled performers. *Frontiers in Psychology*, *6*, 115. https://doi.org/10.3389/fpsyg.2015.00115

Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics*, *36*(6), 565–570. https://doi.org/10.3758/BF03207517

Hedblom, M. M. (2013). The role of working memory in creative insight: Correlation analysis of working memory capacity, creative insight and divergent thinking. https://www.semanticscholar.org/paper/The-Role-of-Working-Memory-in-Creative-Insight-%3A-of-Hedblom/f64332c1608b608c7ef00ee7fa0acbf3f2941669

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. https://doi.org/10.1037/a0027373

Hills, T. T., Todd, P. M., & Jones, M. N. (2015). Foraging in semantic fields: How we search through memory. *Topics in Cognitive Science*, *7*(3), 513–534. https://doi.org/10.1111/tops.12151

Hommel, B. (2015). Between persistence and flexibility. In *Advances in Motivation Science*, *2*, 33–67. https://doi.org/10.1016/bs.adms.2015.04.003

Houpt, J. W., Blaha, L. M., McIntire, J. P., Havig, P. R., & Townsend, J. T. (2014). Systems factorial technology with R. *Behavior Research Methods*, *46*(2), 307–330. https://doi.org/10.3758/s13428-013-0377-3

Houpt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*(5), 341–355. https://doi.org/10.1016/j.jmp.2012.05.004

Howard, Z. L., Belevski, B., Eidels, A., & Dennis, S. (2020). What do cows drink? A systems factorial technology account of processing architecture in memory intersection problems. *Cognition*, *202*, 104294. https://doi.org/10.1016/j.cognition.2020.104294

Huang, P.-S., & Chen, H.-C. (2012). *A version of Chinese remote associates test*. Chinese Behavioural Science Corporation.

Huang, P.-S., Chen, H.-C., & Liu, C.-H. (2012). The development of Chinese word remote associates test for college students. *Psychological Testing*, *58*(4), 581–607.

Innes-Ker, A. H., & Townsend, J. T. (2003, November). *Gestalt perception of emotional expressions. Paper presented at the 44th Annual Meeting of the Psychonomic Society, Vancouver, BC, Canada.*

Jaarsveld, S., & Lachmann, T. (2017). Intelligence and Creativity in problem solving: The importance of test features in cognition research. *Frontiers in Psychology*, *8*, 134. https://doi.org/10.3389/fpsyg.2017.00134

Jones, G. V. (1976). A fragmentation hypothesis of memory: Cued recall of pictures and of sequential position. *Journal of Experimental Psychology: General*, *105*(3), 277–293. https://doi.org/10.1037/0096-3445.105.3.277

Lederman, S. J., Thorne, G., & Jones, B. (1986). Perception of texture by vision and touch: Multidimensionality and intersensory integration. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(2), 169–180. https://doi.org/10.1037/0096-1523.12.2.169

Little, D. R., Altieri, N., Fifić, M., & Yang, C.-T. (Eds.). (2017). *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms*. Elsevier Academic Press. https://doi.org/10.1016/C2015-0-00849-8

Massaro, D. W., Weldon, M. S., & Kitzis, S. N. (1991). Integration of orthographic and semantic information in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 277–287. https://doi.org/10.1037/0278-7393.17.2.277

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 and Social Psychology*, *69*, 511–515. https://doi.org/10.1037/h0041086

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

Mednick, S. A. (1968). Remote associates test. *The Journal of Creative Behavior*, *2*(3), 213–214. https://doi.org/10.1002/j.2162-6057.1968.tb00104.x

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Nielsen, B., Pickett, C., & Simonton, D. (2008). Conceptual versus experimental creativity: which works best on convergent and divergent thinking tasks. *Psychology of Awsthetics, Creativity, and the Arts*, *2*(3), 131–138. https://doi.org/10.1037/1931-3896.2.3.131

Otto, T. U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Current Biology*, *22*(15), 1391–1396. https://doi.org/10.1016/j.cub.2012.05.031

Parker, D. (2019). Cue combinatorics in memory retrieval for anaphora. *Cognitive Science*, *43*(3), e12715. https://doi.org/10.1111/cogs.12715

Preacher, K. J., Rucker, D. D., MacCallum, R. C., & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods*, *10*(2), 178–192. https://doi.org/10.1037/1082-989X.10.2.178

Rubin, D. C., & Wallace, W. T. (1989). Rhyme and reason: Analyses of dual retrieval cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 698–709. https://doi.org/10.1037/0278-7393.15.4.698

Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97–133). MIT Press.

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the remote associates test. *Cognition*, *128*(1), 64–75. https://doi.org/10.1016/j.cognition.2013.03.001

Song, G. W., He, W. G., & Kong, W. (2011). Influence of problem representation and working memory span on pupils' mathematical problem solving.

*Acta Psychologica Sinica*, *43*(11), 1283–1292. http://journal.psych.ac.cn/acps/EN/Y2011/V43/I11/1283

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.

Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review*, *18*(4), 659–681. https://doi.org/10.3758/s13423-011-0106-9

Townsend, J. T., & Wenger, M. J. (2004). A Theory of Interactive Parallel Processing: New Capacity Measures and Predictions for a Response Time Inequality Series. *Psychological Review*, *111*(4), 1003–1035. https://doi.org/10.1037/0033-295X.111.4.1003

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359. https://doi.org/10.1006/jmps.1995.1033

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Wenger, M. J. (1999). On the whats and hows of retrieval in the acquisition of a simple skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1137–1160. https://doi.org/10.1037/0278-7393.25.5.1137

Wenger, M. J., & Townsend, J. T. (Eds.). (2001). Faces as Gestalt stimuli: Process characteristics. *Scientific psychology series. Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 229–284). Erlbaum Publishers.

Wiles, J., Humphreys, M. S., Bain, J. D., & Dennis, S. (1991). Direct memory access using two cues: Finding the intersection of sets in a connectionist model. In R. P. Lippman, J. E. Moody, & D. S. Touretzky (Eds.), *Advances in neural information processing systems* (Vol. 3, pp. 635–641). Morgan.

Yang, C. T., Altieri, N., & Little, D. R. (2018). An examination of parallel versus coactive processing accounts of redundant-target audiovisual signal processing. *Journal of Mathematical Psychology*, *82*, 138–158. https://doi.org/10.1016/j.jmp.2017.09.003

Yang, C.-T., Hsieh, S., Hsieh, C.-J., Fifić, M., Yu, Y.-T., & Wang, C.-H. (2019). An examination of age-related differences in attentional control by systems factorial technology. *Journal of Mathematical Psychology*, *92*, 102280. https://doi.org/10.1016/j.jmp.2019.102280

# Appendix A

## Capacity Bounds

### AND Capacity

The conventional form of Colonius-Vorberg bounds (or C-V bounds; Colonius & Vorberg, 1994) were transformed onto the capacity coefficient space following the inequality developed by Townsend and Eidels (2011):

$$\frac{\log[F_L(t) \times F_R(t)]}{\log[F_L(t) + F_R(t) \ - 1]} \leq C_{AND}(t) \leq \frac{\log[F_L(t) \times F_R(t)]}{\log\{\min[F_L(t),\, F_R(t)]\}}.$$

in which $F_{L,R}$(t) denotes the CDF of the redundant-cue trials, while $F_L$(t) and $F_R$(t) denote the CDF of the two single-cue trials – Left and Right, respectively (Colonius & Vorberg, 1994; Townsend & Eidels, 2011). These bounds can be thought of as upper and lower limits on the unlimited-capacity independent parallel (UCIP) performance (see Little et al., 2017; pp. 71–75).

### OR Capacity

The conventional form of the upper and lower bounds, the race-model bound (Miller, 1982) and Grice bound (Grice et al., 1984); were transformed onto the capacity coefficient space following the inequality rewritten by Townsend and Eidels (2011):

$$\frac{\log\{max[S_l(t),\, S_R(t)]\}}{\log[S_l(t) \times S_R(t)]} \leq C_{OR}(t) \leq \frac{\log[S_l(t) + S_R(t) - 1]}{\log[S_l(t) \times S_R(t)]},$$

where S(t) = 1 − F(t).

(*Appendices continue*)

## Appendix B

## Word List Used in the Word Production Task

**Table B1**
*Word List*

| Exteriors | *M* | *SD* | States | *M* | *SD* | Textures | *M* | *SD* |
|---|---|---|---|---|---|---|---|---|
| 紅色的 red | 5.00 | 0.00 | 圓形的 circular | 4.91 | 0.29 | 玻璃材質的 glass material | 5.00 | 0.00 |
| 橙色的 orange | 5.00 | 0.00 | 橢圓形的 elliptical | 4.86 | 0.35 | 竹製的 bamboo | 5.00 | 0.00 |
| 黃色的 yellow | 5.00 | 0.00 | 正方形的 square | 4.82 | 0.39 | 木質的 woody | 4.95 | 0.21 |
| 綠色的 green | 5.00 | 0.00 | 球形的 spherical | 4.82 | 0.59 | 紙質的 paper | 4.91 | 0.29 |
| 藍色的 blue | 5.00 | 0.00 | 三角形的 triangular | 4.82 | 0.39 | 棉質的 cottony | 4.91 | 0.29 |
| 黑色的 black | 5.00 | 0.00 | 菱形的 rhombic | 4.82 | 0.39 | 塑膠材質的 plastic | 4.86 | 0.47 |
| 白色的 white | 5.00 | 0.00 | 長方形的 rectangular | 4.77 | 0.53 | 石製的 stone | 4.86 | 0.35 |
| 灰色的 gray | 4.95 | 0.21 | 梯形的 trapezoidal | 4.73 | 0.55 | 陶瓷材質的 ceramic | 4.86 | 0.35 |
| 棕色的 light brown | 4.91 | 0.29 | 圓柱形的cylindrical | 4.68 | 0.65 | 布質的 cloth | 4.82 | 0.39 |
| 咖啡色的 dark brown | 4.82 | 0.50 | 錐形的 cone-shaped | 4.50 | 0.60 | 光滑的 smooth | 4.77 | 0.43 |
| 深色的 dark color | 4.73 | 0.63 | 柱狀的 columnar | 4.64 | 0.66 | 粗糙的 coarse | 4.59 | 0.73 |
| 淺色的 light color | 4.36 | 0.95 | 星形的 star-shaped | 4.64 | 0.66 | 金屬材質的 metalized | 4.45 | 0.80 |
| 彩色的 colorful | 4.05 | 0.63 | 倒三角形的 inverted triangular | 4.59 | 0.59 | 毛絨絨的的 fluffy | 4.23 | 1.02 |

**Table B2**
*Word List*

| Exteriors | *M* | *SD* | States | *M* | *SD* | Tastes | *M* | *SD* |
|---|---|---|---|---|---|---|---|---|
| 胖的 Fat | 4.77 | 0.53 | 氣態的 gaseous | 5.00 | 0.00 | 酸味的 acid | 5.00 | 0.00 |
| 長的 Long | 4.73 | 0.55 | 液態的 liquid | 5.00 | 0.00 | 甜味的 sweet | 5.00 | 0.00 |
| 短的 Short | 4.73 | 0.55 | 固體的 solid | 4.91 | 0.43 | 苦味的 bitter | 5.00 | 0.00 |
| 高的High | 4.73 | 0.55 | 冰的 icy | 4.86 | 0.35 | 辛辣的 spicy | 5.00 | 0.00 |
| 矮的 Dwarf | 4.73 | 0.55 | 燙的 hot | 4.77 | 0.43 | 鹹味的 salty | 5.00 | 0.00 |
| 大的 Big | 4.68 | 0.65 | 熱的 thermal | 4.73 | 0.46 | 酒味的 alcohol | 4.77 | 0.43 |
| 小的 Small | 4.68 | 0.65 | 冷的 cold | 4.73 | 0.46 | 臭味的 smelly | 4.68 | 0.72 |
| 瘦的 Thin | 4.68 | 0.57 | 僵硬的 hard | 4.68 | 0.78 | 沙茶味的 barbecue | 4.68 | 0.57 |
| 扁的 Flat | 4.68 | 0.65 | 柔軟的soft | 4.64 | 0.79 | 薄荷味的 mint | 4.64 | 0.85 |
| 薄的 Filmy | 4.59 | 0.73 | 發光的 luminous | 4.59 | 0.59 | 芳香的 fragrant | 4.50 | 0.91 |
| 厚的 Thick | 4.55 | 0.74 | 稀疏的 sparse | 4.45 | 0.80 | 油膩的 greasy | 4.45 | 0.80 |
| 寬廣的 Vast | 4.36 | 0.85 | 溫暖的 warm | 4.45 | 0.80 | 腥味的 fishy | 4.36 | 0.85 |
| 狹窄的 narrow | 4.36 | 0.90 | 有彈性的 elastic | 4.45 | 0.91 | 清淡的 mild | 4.32 | 0.78 |

(*Appendices continue*)

## Appendix C

## 2-Character Chinese Words Used in the OSPAN Task

**Table C1**

| Presentation condition | Trial 1 | Trial 2 | Trial 3 |
|---|---|---|---|
| 2 Presentations | 飛機 aero plane | 果汁 juice | 鏡子 mirror |
| | 名片 business card | 蝴蝶 butterfly | 台南 Tainan |
| 3 Presentations | 外套 coat | 汽車 cat | 冰箱 refrigerator |
| | 漢堡 hamburger | 口紅 lipstick | 練習 practice |
| | 課本 textbook | 唱片 phonorecord | 蕃茄 tomato |
| 4 Presentations | 照片 photograph | 滑鼠 mouse | 面紙 tissues |
| | 零件 machine parts | 窗簾 curtain | 燈泡 lamp bulb |
| | 劇本play script | 大象 elephant | 樹葉 leaf |
| | 泥土 muddy soil | 字典 dictionary | 指甲 nail |
| 5 Presentations | 石頭 stone | 數學 mathematics | 帽子 hat |
| | 日曆 calendar | 蠟燭 candle | 鋼琴 piano |
| | 玩具 toy | 員工 employee | 卡通 cartoon |
| | 海洋 ocean | 操場 playground | 案例 case |
| | 茶壺 teapot | 靈感 inspiration | 沙漠 desert |
| 6 Presentations | 友誼 friendship | 分開 separation | 強迫 force |
| | 檔案 archive | 香水 perfume | 公寓 apartment |
| | 天空 sky | 過敏 allergy | 眉毛 eyebrow |
| | 水餃 dumpling | 鋼筆 pen | 馬路 road |
| | 演奏 play | 美國 America | 報紙 newspaper |
| | 皮夾 wallet | 杯子 cup | 企鵝 penguin |

Received February 27, 2020
Revision received December 25, 2020
Accepted December 31, 2020 ■