BOUNDARY CONDITIONS

FOR A VISUAL WORKING MEMORY CAPACITY MODEL

A Dissertation presented

to the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment of the Requirements

for the Degree Doctor of Philosophy

---

by

Zhijian Chen

Dr. Nelson Cowan, Dissertation Supervisor

JULY 2009

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation titled

BOUNDARY CONDITIONS FOR A VISUAL WORKING MEMORY CAPACITY MODEL

Presented by Zhijian Chen,

a candidate for the degree Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Nelson Cowan, Committee Chair

Professor Moshe Naveh-Benjamin, Committee Member

Professor Jeffery Rouder, Committee Member

Professor John Kerns, Committee Member

Professor Judith Goodman, Committee Member

ACKNOWLEDGEMENTS

I would like to express my gratitude to my dissertation advisor, Dr. Nelson Cowan, who gave me the opportunity to pursue a Ph.D. degree in cognitive psychology and provided me with financial support and academic guidance throughout my graduate studies. I am deeply indebted for the benefits to my professional career from the treasures I found in his wisdom, knowledge and kindness.

I am especially grateful to my wife, Jiemei (Stella) Ruan, whose patient love enabled me to complete my Ph.D. program and dissertation over the years.

This project was also supported by a MU Department of Psychological Sciences Dissertation Research Grant awarded to the author.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Cowan's K (Cowan, 2001) is an effective measure of visual working memory capacity and has been widely accepted by working memory researchers in recent years. However, one fundamental assumption of this measure, so called the all-or-none assumption (i.e., an observer either remembers everything or nothing about a visual stimulus), could be over-simplified and thus impose validity boundaries for Cowan's K. Little has been done to examine the potential limitations of the assumption in terms of its impact on Cowan's K. As an attempt to fill this gap in the literature, the present project explored the boundary conditions of Cowan's K through addressing the following questions: 1) whether the all-or-none assumption can survive different types of visual working memory test procedures and 2) whether the all-or-none assumption is still viable with complex visual stimuli. The results suggest that the all-or-none assumption is fairly robust across different test procedures, as long as the visual stimuli are simple and familiar. When complex and novel stimuli were used, however, the all-or-none assumption led to significant discrepancies across different test procedures. In addition, it was found that the validity of the all-or-none assumption could be compromised (when articulatory suppression was imposed during visual presentation of English consonants) or enhanced (when extra time was given during visual presentation of Chinese characters). In conclusion, the all-or-none assumption can be extended to visual working memory test procedures other than the conventional change detection paradigm, provided that features of a visual object can be encoded into an abstract representation in visual working memory. Reasons for the failure of the all-or-none assumption were discussed and a revision of Cowan's K formulation was proposed to account for visual working

memory performance with both simple and complex stimuli. An interesting observation of experiment participants' non-rational strategy use was also reported.

Introduction

Concept of Working Memory Capacity

People can attend to only a tiny portion of an ocean of information available from the outside world or past experiences. Information in the focus of attention is highly accessible in various cognitive activities but severely limited in amount (Atkinson & Shiffrin, 1968; Broadbent, 1958, 1975; Cowan, 1995; James, 1890; Miller, 1956). The cognitive mechanism for holding and processing this limited amount of transient information is known as working memory (WM) (Baddeley, 1986; Baddeley & Hitch, 1974; Cowan, 2001; Daneman & Carpenter, 1980; Miller, Galanter, & Pribram, 1960; Miyake & Shah, 1999). Assessment of working memory capacity has remained one of the central topics in cognitive psychology since its establishment in 1950s, not only because of its theoretical significance but also its important psychometric applications, as working memory capacity has exhibited excellent predictive power for individual differences in higher-level cognition (Conway, Kane, & Engle, 2003; Cowan, 2001; Cowan, 2005; Hannon & Daneman, 2006).

In particular, WM capacity for visual information has been intensively studied in recent years (Awh, Barton, & Vogel, 2007; Bays & Husain, 2008; Luck & Vogel, 1997; Rouder et al., 2008; Song & Jiang, 2006; Vogel, Woodman, & Luck, 2001; Zhang & Luck, 2008), with behavioral, electrophysiological and neural imaging data consistently implying that the norm of visual working memory (VWM) capacity might be as small as 3~4 objects (Cowan, 2001; Luck

& Vogel, 1997; McCollough, Machizawa, & Vogel, 2007; Todd & Marois, 2004), although there are disagreements over whether the unit of VWM capacity should be an object, an object part/feature conjunction or the information load of an object (Bays & Husain, 2008; Davis & Holmes, 2005; Sakai & Inui, 2002; Wilken & Ma, 2004).

Measurement of Visual Working Memory Capacity

The standard measurement of individual VWM capacity involves a two-stage procedure: 1) collection of behavioral data about how well an observer can notice a change between two visual arrays over a brief period of time with a procedure called the change detection paradigm (Luck, 2007; Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Vogel et al., 2001) and 2) estimation of VWM capacity with a mathematical model based on the behavioral accuracy data (Palmer, Verghese, & Pavel, 2000; Pashler, 1988; Rouder et al., 2008).

A typical version of the change detection paradigm uses simple objects such as color squares as stimuli (e.g. Luck & Vogel, 1997; Morey & Cowan, 2004; Rouder et al., 2008). At study, the observer is presented with a sample array of colored squares for a few hundred milliseconds. At test, an array is shown and remains until the observer makes a response. The test array is either identical to the sample array or differs in a critical item. The observer is required to determine whether the critical item has changed or not based on his/her VWM of the item in the sample array. A change detection test consists of a number of “change” and “same” trials. In a “change” trial, the sample and test arrays differ in a critical item. In contrast, if the critical item doesn’t change between the sample and test arrays, this is called a “same” trial. As the participant has to make a forced decision between “change” and “same”, four combinations of trial type and response type are possible: 1) a hit (i.e., a “change” response in a “change” trial), 2) a miss (i.e., a “same” response in a “change” trial), 3) a correct rejection (i.e., a “same” response in a “same” trial) and 4) a false alarm (i.e., a “change” response in a “same” trial). Over the entire set of “change” trials, the proportion of hits is known as the hit rate and similarly, over the entire set of “same” trials, the proportion of correct rejections is known as the correct rejection rate. It is obvious that the miss rate would be (1 – hit rate) and the false alarm rate would be (1- correct rejection rate).

A variant of the typical change detection paradigm uses a single-item or encircled-item probe rather than a whole array at test for the change-or-same judgment. In this method the participant is informed as to which item may have changed and only needs to make the “change” versus “same” decision. One major advantage of the single item or encircled-item probe is that it reduces the attention cost of perceiving and processing multiple test items. Due to the shared mechanism between visual memory and visual perception, the single-item probe approach can minimize the interference between viewing the test array and holding the sample array in VWM (Wheeler & Treisman, 2002).

Theoretically, if an observer’s VWM capacity is big enough to contain all information of the sample array, the critical item should be in VWM as well, so that his/her hit rate should be as high as “1” and false alarm rate as low as “0”. However, if his/her VWM capacity is smaller than what it takes to contain the entire sample array, the critical item could fall out of VWM from time to time. The more items in the sample array, the higher the chance that the critical item is not in VWM. With the critical item not in VWM, a hit/correct rejection is still possible but only by guessing. The extent of guessing can be ascertained from the relation between the hit rate and the false alarm rate. In general, the more the hit rate exceeds the false alarm rate, the less guessing was used in the task. An estimate of VWM capacity has to take into account the contamination from guessing in the behavioral data.

All-or-none Assumption of Cowan's K

A popular VWM capacity estimate, known as Cowan's K, can be arithmetically calculated from array-comparison data using a double high threshold (DHT) model with the hit and false alarm rates from a single-item-probe change detection test (Cowan, 2001). A key simplifying assumption underlying the model is that the participant's knowledge of each item is all-or-none; the participant is assumed not to have partial knowledge of any stimulus. Cowan's K has been widely adopted in VWM studies (Kawasaki, Watanabe, Okuda, Sakagami, & Aihara, 2008; Olson, Moore, & Drowos, 2008; Saults & Cowan, 2007; Song & Jiang, 2006; Todd & Marois, 2004) as a simple and expedient measure. Moreover, Cowan's K proves a fairly robust estimate of VWM capacity even under rigorous quantitative assessment (Cowan & Rouder, 2009; Rouder et al., 2008).

The DHT model assumes that 1) a discrete representation of a stimulus in the sample array is either in or out of VWM, and 2) when the critical item in the sample array is held in VWM, the probability of a hit or a correct rejection is 1; otherwise, the response is made on the basis of some rate of guessing “change” versus “same”. This set of premises about VWM is termed as the all-or-none assumption in the present paper. Furthermore, Cowan (2001) posits that given a sample array of *N* objects, *K* of them can be entered in to VWM. The probe item is randomly selected, so that the probability of a probed stimulus being held in VWM capacity is Pr(a probed stimulus in VWM) = K/N. Accordingly, the DHT model can be characterized by equations (1) ~ (4).

When $K \leq N$,

$$ht = K/N+(1-K/N)*g \tag{1}$$

$$cr = K/N+(1-K/N)*(1-g) \tag{2}$$

where ht = Pr (hit), cr = Pr (correct rejection), g = Pr(guessing of a “change”), K = the size of the discrete VWM capacity, and N = the array size.

When $K > N$,

$$ht=1 \tag{3}$$

$$cr=1 \tag{4}$$

A formula of the VWM capacity estimate *K* derived from the two equations is as follows.

$$K=N*(ht+cr-1) \tag{5}$$

Assuming the performance in each trial is independent and identically distributed, the frequencies of hits and correct rejections are distributed as conditionally independent binomials (Rouder et al., 2008). Therefore, an individual’s VWM capacity can be conveniently calculated by substituting the observed hit rate, correct rejection rate and the corresponding array size from a change detection test into equation (5). Although Cowan's K is considered to be the first approximation of VWM capacity, a number of VWM studies have shown that it serves the purpose fairly well, exhibiting some good qualities of a VWM capacity measure such as 1) asymptoting at a constant of ~3-4 over increasing array sizes (Cowan, 2001), 2) being consistent with other neural imaging and psychophysiological indexes of VWM capacity (Todd & Marois, 2004; Vogel & Machlzawa, 2004) and 3) surviving selective influence on the guessing biases (Rouder et al., 2008).

Nevertheless, little has been done to explore the boundary conditions for Cowan's K. Theoretically, the all-or-none assumption underlying Cowan's K should be independent of specific VWM capacity measurement procedures, provided that the formula is adjusted to match the testing situation. However, so far, it has been tested only with the change detection paradigm. Can it be extended to other procedures and yield a converging VWM capacity estimate with Cowan's K?

In addition, the all-or-none assumption has been validated mostly with simple objects that are made up of one or two familiar features such as color and regular shape (Luck & Vogel, 1997; Rouder et al., 2008; Zhang & Luck, 2008). Given that recent studies have shown that complexity of visual objects may affect VWM representations (Alvarez & Cavanagh, 2004; Awh et al., 2007; Eng, Chen, & Jiang, 2005; Wheeler & Treisman, 2002), can the all-or-none assumption still stand with complex visual stimuli?

We know that performance decreases when complex objects are used as stimuli but there are alternative ways to explain it. Alvarez and Cavanagh (2004) used a change detection procedure to measure the VWM capacity and a visual search task to assess the degree of complexity for 5 stimulus types, respectively. They found that the measured VWM capacity ranged from a high of 4.4 items for the simplest stimuli, colors, to a low of 1.6 items for the most complex stimuli, shaded cubes of different orientations. The actual reason for the decrease with complexity, though, could occur for two reasons: because the number of items in working memory depends on complexity, or because the resolution of each item in memory depends on complexity. The latter possibility violates the all-or-none assumption.

Subsequent studies suggest that the all-or-none assumption is violated for complex items. Awh et al. (2007) replicated Alvarez and Cavanagh (2004)'s findings, but they also found that the apparent VWM capacity reduction was associated with the increased similarity between the sample and test arrays. For example, a cube in a certain orientation changing to a colored spot as a probe was as easy to detect as a color changing to a different color, and much easier than a cube changing to a differently oriented cube. Therefore, they suggested that visual complexity might impair the resolution of VWM representations so that the decreased performance in the change detection procedure with increased complexity in stimuli (Alvarez & Cavanagh, 2004) was presumably caused by the difficulty in comparing items in the sample and test arrays, rather than by fewer items in memory.

The present work tests another implication of the findings of Alvarez and Cavanagh (2004) and Awh et al. (2007). Assume that the effect of using complex items is to lower the capacity, contrary to what these authors conclude, and that the all-or-none assumption still holds. Assume also that there is a convergence between estimates of capacity (*K*) obtained for simple objects using different test procedures. Then there should be a convergence of capacity estimates also for complex items, but at a lower capacity level. If, on the other hand, using complex items lowers the resolution but not the capacity (i.e., the all-or-none assumption is violated), then the convergence might be lost.

Moreover, further thoughts about the nature of the lowered resolution might allow an orderly analysis of the results. For example, suppose that each complex stimulus such as a Chinese character contains, for individuals naïve to that language, at least two features. It might be that not all features can always be integrated into a single representation in VWM, a violation of the all-or-none assumption. A correct response may not be made on the basis of only one of the features of the probed character that is partially represented in VWM. In that case, a revised capacity analysis could be carried out in order to observe convergence between VWM measures of different procedures.

Quest of Boundary Conditions of the All-or-none Assumption

The present project was primarily to investigate the boundary conditions of the all-or-none assumption underlying Cowan’s K. To address the two questions just raised, a total of 4 experiments was carried out. Experiment 1a was designed to directly address the first question that whether the all-or-none assumption can be generalized beyond the change detection paradigm. The answer turned out to be positive, for English consonants (though the model that worked entailed a non-rational use of information). Experiment 1b examined the second question whether the all-or-none assumption holds for VWM capacity tests with complex objects; and the answer appears to be “no” and the model was extended to the case of multiple features per item. Experiment 2-3 were follow-up studies to replicate the findings of Experiment 1a & 1b and also to shed light on some unsettled issues.

In Experiment 1a, the response procedure of a single-item probe change detection task was modified to create another two VWM tasks, namely, recalling a critical item’s identity or location from multiple choices. A schematic illustration of all three tasks is shown in Figure 1. In general, each trial in the experiment consisted of a sample array display, a delay interval, a mask display, and a test array display. The multiple-choice identity/location recall tasks differed from the change detection task only in the test array display. For the location recall task, an identity probe (i.e., “B”) was provided and the participant was asked to point out the location of the cue from four location options. Similarly, for the identity recall task, a location probe (i.e., “?”) was provided and the participant was asked to choose the identity of the probe from four options.

The rationale behind the multiple-choice identity/location recall task as a test of VWM capacity will be set out below, but I will describe two possible models. One model is fully rational and makes full use of the available information in VWM, whereas the second model (which turned out to produce much better convergence between procedures) is not fully rational because it makes only partial use of the information in VWM. According to the all-or-none assumption, if a critical item is in VWM, the observer should be able to hold both its identity and location features in mind and certainly make a correct choice from the options. The two possibilities pertain to what happens if the critical item is not in VWM. In that circumstance it is possible that (1) the observer can make full use of information about the non-critical items available in VWM; or (2) cannot take advantage of information of non-critical items in VWM.

The issue is as follows. If one is shown a probe (e.g., an item) and N response choices (e.g., locations), what happens when the correct response choice is not in VWM? Can some of the other response choices be ruled out on the basis of other items that *are* in VWM? With Possibility 1, let K_1 denote VWM capacity under the assumption of full use of information. The probability that the probed item is in VWM at the time of test is K_1/N. When it is not in VWM, which happens with probability $1- K_1/N$, the chance of guessing correctly is then $1/(N-K_1)$, because participants could first eliminate K_1 non-critical items they held in VWM capacity from N options and then make a random guessing among $(N- K_1)$ options. Therefore, with a similar logic to the formula of Cowan's K, the probability of an accurate response in multiple-choice recall paradigms can be characterized by equation (6).

When $K_1 < N$,

$$Ac= K_1/N+(1-K_1/N)*(1/(N- K_1)) \qquad (6)$$

where Ac = Pr (Accurate response), K_1= the size of discrete VWM capacity in a formula assuming full use of information, and N = the array size = the number of options;

When $K_1 \geq N$,

$$Ac=1 \quad (7)$$

Some algebraic transformations of equation (6) give rise to a formula of K_1, assuming full use of information as follows.

$$K_1=N*(Ac-1/N) \quad (8)$$

With Possibility 2, the observer is not able to take advantage of any information about non-critical items and thus, when the probed item is not in VWM, makes one random choice out of N options (instead of $N- K_1$ options as before). That is, if the probed item is not in VWM, the participant simply guesses among all the items. In this case the probability of a correct guess is $1/N$. Let K_2 denote VWM capacity without assuming full use of information. The probability of an accurate response in multiple-choice recall paradigms can be characterized by equation (9).

When $K_2 < N$,

$$Ac= K_2/N+(1- K_2/N)*(1/N) \quad (9)$$

where Ac = Pr (Accurate response), K_2= the size of the discrete VWM capacity in a formula assuming no use of information about non-critical item in VWM , and N = the array size = the number of options;

When $K_2 \geq N$,

$$Ac=1 \quad (10)$$

Using equation (9), K_2 can be expressed as a function of *Ac*, which is formulated as equation (11) after some algebraic transformation.

$$K_2 = N*(Ac-1/N)*(N/(N-1)) \qquad (11)$$

Therefore, two different ways to estimate VWM capacity are to calculate K_1 or K_2 from equation (8) or (11), respectively.

If the all-or-none assumption is valid, Cowan's K from the change-detection task should converge with either K_1 or K_2 from the remaining tasks, if one of these models holds true. To anticipate the results, Experiment 1a suggested that the all-or-none assumption fits data, but only with Possibility 2, no use of VWM for non-critical items. In the subsequent experiments, any derivative VWM capacity estimates are therefore calculated using the K_2 formula unless it is otherwise stated.

In Experiment 1b, I examined whether Cowan's K is consistent with alternative VWM capacity measures when complex and novel visual stimuli (i.e. Chinese characters for observers who did not speak Chinese) were used. The results were compared to Experiment 1a, in which simple and familiar visual stimuli (i.e. English consonants for observers who spoke fluent English) were used. The rationale behind the comparison is as follows. If the representation of the critical item is indeed all-or-none in VWM, the three paradigms with different response modes should still return the same VWM capacity estimates regardless of stimulus complexity as long as the contamination from different guessing processes across three paradigms has been quantitatively accounted for. Consequently, if the all-or-none assumption remains valid for both simple and complex objects, VWM capacity values estimated from different procedures should make no difference for the same type of stimuli and thus converge to each other, even though complex stimuli may result in lower VWM capacity estimates. To anticipate, that was not the case but an alternative model that is not all-or-none did work.

Common Method

In general, the core paradigm of all 4 experiments reported in the present paper was in the form of three VWM tasks, each with a distinctive response mode. Each trial of the tasks consisted of a sample array display, a delay interval, a mask display, and a test array display. In each experiment, the response mode required for a specific trial was unknown until the test array display was shown. In this section, the common parameters or characteristics of all 4 experiments in the project will be delineated below. The variation from the common method in a specific experiment will be indicated when that experiment is described in details.

Participants. A total of 84 undergraduates (age 18-25 years old) from introductory psychology courses at University of Missouri, Columbia participated in the project in exchange for course credit.

Participants naïve to the task were recruited through an on-line experiment management system of the Department of Psychological Sciences of the University of Missouri at Columbia. Participants read and signed an informed consent form before the experiment started. All participants were fluent speakers of English with normal hearing and normal or corrected-to-normal vision. No participants in an experiment with Chinese characters as stimuli could read Chinese.

Apparatus & Stimuli. VWM tasks in this project were computerized with E-prime software (Schneider, Eschman, & Zuccolotto, 2002) and administered on personal computers equipped with 15-inch CRT monitors, loudspeakers and a mouse-pointing device in quiet, private booths.

Two types of visual objects were used as stimuli in the VWM tasks. One type was English consonants with low phonological and visual confusability, with the entire stimulus set comprising {*B, G, H, K, N, J, T*}. Presumably, for participants who speak English fluently, these stimuli could be easily verbalized and conceptualized but are unlikely to be formed into words due to the absence of vowels in the array. The other type was Chinese characters, with the entire stimulus set comprising {刀, 人, 上, 七, 又, 厶, 匚}. Each of these Chinese characters had no more than 3 strokes and was structurally comparable to the English consonants but, for participants who did not know Chinese, it was rather difficult to conceptually and verbally encode these characters.

As shown in Figure 1, when the sample and test array were displayed, the computer screen area was marked into two parts by a horizontal white line. The upper part (12 x 7.5 inch), called the stimulus area, was designated for presentation of stimuli and the lower part (12 x 1.5 inch), called the response area, for response options. When presented in the stimulus area, each English consonant or Chinese character was shown in black and filled up a 1 x 1 inch white box (3° x 3° visual angle, with an assumed viewing distance of about 20 inches). Stimuli were displayed on a gray background and pseudo-randomly distributed within a 10.5 x 6 inch area (30° x 17° visual angle) that was centered in the stimulus area. The minimal side-to-side distance among stimuli was about 2° of visual angle. Each quadrant of the stimulus area had no more than two stimuli at any time. In a sample array display, the response area below the white line was always blank, but in the test array display, the response area showcased stimulus items or the “Change” and “Same” buttons that were evenly aligned with 1-inch side-to-side separation from each other.

The SPACE BAR had to be pressed to initiate each trial by the participants. The participants used the computer mouse to click on one of the designated options or buttons on the screen to make a response.

Procedure. After participants returned the signed informed consent form, they were situated in a quiet and private booth, facing a computer screen in a distance of about 20 inches. The experimenter then launched the program of the experiment and completed a series of survey questions on the computer by filling in the participant’s subject number and demographic information.

Following the survey session was a practice trial session, in which the experimenter first explained the instructions with a schematic illustration of a task and then had the participant carry out a couple of trials for practice. There were three tasks in each experiment. All participants appeared to have learned the task after two practice trials. After the practice trial session, the participant was left alone in the booth and completed the formal test session at his or her own pace.

The scores of the practice trials were excluded from the data analyses. The formal test scores from all participants who successfully finished the experiment were included in the analyses, as there was no evidence of systematic problems or mistakes in the data and no one performed at or below chance level over all conditions in an experiment.

For all four experiments in the present project, the major statistics of interest were mean proportion correct in the VWM tasks and VWM capacity estimates, namely Cowan's K in tasks with change detection response mode and two other derivative Ks in tasks with multiple-choice identity or location recall response mode. Cowan's K was calculated according to equation (5), with the probability of hit, *ht*, estimated by the hit rate and the probability of correct rejection, *cr*, estimated by the correct rejection rate. The other two derivative K_1s or K_2s were calculated according to equation (8) or (11), respectively, with the probability of accurate response, *Ac*, estimated by the proportion correct of responses. Tests of statistical significance included ANOVAs and the post-hoc Newman-Keuls tests. The alpha level was set to 5%. For each significant result, an F value, p value, and effect size, partial Eta squared, η_p^2 (Cohen, 1973) will be reported in the present paper. All statistical analyses were implemented with Statistica software (StatSoft, 2008).

Project Overview. A brief preview of the studies and findings is as follows. In Experiment 1a, participants completed three VWM tasks with English consonants as stimuli. There were three response modes, including change-or-same detection, multiple-choice identity recall and multiple-choice location recall modes. Three VWM capacity estimates (i.e., Cowan's K and two derivative K_2s), all based on the all-or-none assumption, yielded remarkably similar results and, across arrays with 3 to 6 items, asymptoted at about 4 items in accordance with previous work. This study shows that the logic of the all-or-none working memory capacity assumption can be generalized across procedures, though not with a fully rational model.

The remaining experiments show, however, that the logic does not generalize to situations in which encoding of the items is difficult. In Experiment 1b, participants completed the same tasks as in Experiment 1a except that the visual stimuli were Chinese characters. However, there was no stable asymptote and the measure of VWM capacity estimates proved to be very low and dependent on the task. The data were accounted for much better with a model in which partial information about each item was a possibility. In Experiment 2, participants conducted three VWM tasks similar to those in Experiment 1a, except that during the maintenance of the array they performed a secondary task in which rehearsal is prevented, i.e., articulatory suppression (Chen & Cowan, 2009). The VWM capacity estimate value was again lower than in Experiment 1a and again diverged across tasks. Finally, in Experiment 3, participants carried out three VWM tasks similar to those in Experiment 1b, except that one half of the trials had a sample array display duration of 500 msec, as before, whereas the other half had a display duration of 1000 msec, longer than before. The data from trials with 500 msec display duration basically replicated the pattern in Experiment 1b. More importantly, confirming the importance of adequate encoding, the longer display duration in the sample array not only boosted all three VWM estimates, but also removed discrepancies among them.

Experiment 1: VWM Tasks with 3 Response Modes and English Consonants (1a) or Chinese Characters (1b) as Stimuli

Methods of Experiment 1a and 1b

Experiment 1a and 1b shared the same method except that they had two different groups of participants and used English consonants and Chinese characters as visual stimuli, respectively.

Participants. There were 15 participants in each of Experiment 1a (9 females, 6 males) and Experiment 1b (11 females, 4 males).

Stimuli and Conditions. In Experiment 1a, the sample array consisted of 3 to 6 English consonants, which were randomly selected without repetition from database {B, G, H, K, N, J, T}. Similarly, in Experiment 1b, 3 to 6 Chinese characters were randomly selected without repetition from database {刀, 人，上，七，又，厶，匸}.

Combination of the array size (i.e., 3,4,5, and 6) and response mode (i.e., change-or-same detection, multiple-choice identity recall, and multiple-choice location recall) factors resulted in a total of 12 conditions. Each condition included 40 trials and 50% of trials in the change-or-same detection condition were “change” trials.

Procedure. The procedure of Experiment 1a & 1b is illustrated in Figure 1. At the beginning of the experiment, the experimenter explained a VWM task with a schematic diagram shown on the computer screen and then supervised the participant to do two practice trials so as to make sure the instruction was followed properly. The practice trials were identical to trials in the formal test session. Experiment 1a and 1b each consisted of three VWM tasks.

The participant initiated the task by pressing “SPACE BAR” when ready. First of all, a fixation sign was shown for 400 msec, signaling where the participant to look for the upcoming sample array. Then came a sample array, which was centered in the stimulus area against a gray background, and lasted for 500 msec. After that, a clean gray screen replaced the entire sample array for 1000 msec and was followed by an array of white boxes which masked exactly over the locations of stimuli in the sample array for 500 msec. This was intended to remove the lingering visual sensory memory of the sample array without introducing unnecessary interference into VWM. After that, a test array was then shown and stayed until the participant made a response with the mouse-pointing device. The white boxes in the test array had the same spatial configuration as the stimuli in the sample array, but they had a unique characteristic for each response mode (See Figure 1), as follows.

In the change-or-same detection mode, all but one of the white boxes were empty and one contained a probe item. In the response area, there were a “change” and a “same” buttons. The probe item was either the same item as the one at that location in the sample array or, in change trials, one from the stimulus set but not in the sample array. The appropriate response button was to be clicked.

In the multiple-choice identity recall mode, all but one of the white boxes were empty and one box contained a question mark. In the response area, all items from the sample array were provided, but in random order. The participant was to mouse-click on a single option matching the item previously placed where the question mark was shown.

In the multiple-choice location recall mode, the white boxes were all empty. A probe item from the sample array was provided at the center of the response area. The participant was to mouse-click on one of the white boxes to indicate where the probe item was previously located in the sample array.

Both visual and auditory feedbacks were provided immediately following the participant's response. If the response was correct, the text message said "good job" and there was a chime; otherwise, a text message said "work harder" and there was a tone. Participants initiated each trial when ready and were allowed to take breaks between trials as long as needed. All trials in an experiment were arranged in a random order.

Results

Experiment 1a. In Experiment 1a, 15 participants completed three VWM tasks with different response modes, including 1) change-or-same detection, 2) multiple-choice identity recall and 3) multiple-choice location recall of English consonants. Each participant contributed 480 accuracy scores to the following analyses.

The initial analysis of Experiment 1a was to decide which one of two possible strategies for multiple-choice recall tasks was realistic in the present project. VWM capacity estimates K_1 and K_2 were calculated and plotted along with Cowan's K in Figure 2 and the left-hand panel of Figure 3, respectively. A one-way ANOVA and its post hoc Newman-Keuls tests showed Cowan's K was strikingly higher than the K_1s from the identity and location recall tasks, with a main effect of response mode ($F(2, 28)=72.08$, $p < .001$, and $\eta_p^2 = .84$) and with two significant post hoc Newman-Keuls comparisons between Cowan's K and both K_1s (p's < .001).

In stark contrast, as the left-hand panel of Figure 3 shows, three VWM capacity estimates, namely Cowan's K and the two derivative K_2s, converged almost perfectly. They started at about 3 for an array size of 3 and rose gradually with increasing array sizes, but became asymptotic at about 4 for array sizes of 5 and 6. This observation was confirmed by a series of statistical tests. There was no effect of response mode, $F(2, 28)=.23$, $p=.798$, $\eta_p^2 = .02$, and there was no interaction between array sizes and response modes, $F(6, 84)=.35$, $p=.911$, $\eta_p^2 = .02$.

In addition, the convergence between Cowan's K (for change detection) and the two K_2 estimates (for identity and location tasks) was confirmed by a series of Bayesian t tests for each array size. These tests are capable of providing an odds ratio between the likelihood of the null and alternative hypotheses (Rouder, Speckman, Sun, Morey, & Iverson, 2009). A statistic termed JZS Bayes Factors, based on a particular, typical alternative hypothesis, showed reasonable plausibility for the null hypothesis in the comparison of Cowan's K and the average K_2 across the identity recall and location recall tasks. When the ratio is above 1.0, it favors the null hypothesis. The obtained null:alternative ratio was 5 for an array size of 3, 5.1 for an array size of 4, 3.8 for an array size of 5, and 3.8 for an array size of 6. For example, for an array size of 3, it was 5 times more probable that the null hypothesis was true than that the specified alternative hypothesis was true. These results yield reasonably good evidence for the convergence between Cowan's K and K_2s.

One-way ANOVAs of capacity across set sizes, separately for each VWM capacity estimate, were significant: for Cowan's K with the change detection mode, $F(3, 42)=11.46$, $p <.001$, $\eta_p^2 = .45$; for K_2 with the multiple-choice identity recall mode, $F(3, 42)=13.52$, $p<.001$, $\eta_p^2 = .49$, and for K_2 with the multiple-choice location recall mode, $F(3, 42)=15.880$, $p<.001$, $\eta_p^2 = .53$. Post hoc Newman-Keuls tests showed significant difference between array sizes 3 and 4 (p's < .01) in all VWM capacity estimates but no significant difference in any of them between array sizes of 5 and 6.

Given that these findings clearly showed that Possibility 2, above, the non-rational model, was the realistic outcome of the present project, the derivative VWM capacity estimates were calculated with the K_2 formula in the present paper unless it is otherwise stated. Descriptive statistics on VWM capacity estimates and accuracy associated with different response modes across array sizes can be found in Table 1 and 2. The mean VWM capacity estimates ranged from 2.92 to 4.16 across all conditions.

An analysis of the proportion correct was carried out to assess response biases. As shown in the upper left-hand panel of Figure 4, "change" trials and "same" trials departed from each other in accuracy further and further away with increasing array sizes. This observation was supported by a significant interaction between array sizes and trial types (i.e, "change" versus "same" trials), F(3,42) =

3.01, p = .04, η_p^2 = .18. Also, post hoc Newman-Keuls tests showed no significant difference between "change" and "same" trials at array sizes of 3, 4 and even 5, but the difference turned significant at the array size of 6, p < .001. In addition, there was a significant main effect of trial type (F (1,14) = 7.30, p < .05, η_p^2 = .34) and a significant main effect of array size (F (3,42) = 27.73, p < .001, η_p^2 = .66). This pattern suggested that bias in guessing toward "change" as opposed to "same" seems to increase with increased array sizes.

Experiment 1b. Experiment 1b was almost identical to Experiment 1a in the method except that Experiment 1b used Chinese characters as stimuli and a different set of 15 participants. Descriptive statistics on VWM capacity estimates and accuracies in Experiment 1b are also reported in Table 1 and 2. The mean VWM capacity estimates ranged from 1.18 to 2.19 across all conditions. Despite the similarities in procedure between experiment 1a and 1b, there were dramatically different patterns in the results of the two experiments, presumably due to their difference in the stimulus type.

As the right-hand panel of Figure 3 shows, the three VWM capacity estimates appeared to first decline and then asymptote with increasing array sizes. One-way ANOVAs showed that both derivative K_2s calculated based on the multiple-choice recall trials had a significant main effect of array size, for identity recall, F(3, 42)=4.23, p=.011, η_p^2 = .23, for location recall, F(3, 42)=4.28, p=.010, η_p^2 = .23. Moreover, post hoc Newman-Keuls tests showed that K_2s were always higher at an array size of 3 compared to at an array size of 6 (p's <.05) and there were no significant differences between array sizes 5 and 6. Nevertheless, Cowan's K calculated from the change detection trials had no significant array size effect.

In contrast to Experiment 1a, one striking effect observed from Experiment 1b was that Cowan's K appeared inconsistent with the other two derivative K_2s. A two-way (3 response modes X 4 array sizes) ANOVA showed that there was a significant main effect of response mode, $F(2, 28)=7.90$, $p < .01$, $\eta_p^2 = .36$ and a post hoc Newman-Keuls test confirmed that Cowan's K was significantly lower than the other two derivative Ks (p's < .01), whereas the two derivative K_2s didn't differ from each other. In addition, a significant main effect of array size ($F(3, 42)=6.55$, $p<.001$, $\eta_p^2 = .32$) but no significant interaction between response mode and array size were found.

Another remarkable difference between Experiment 1b and Experiment 1a was in the pattern of accuracy between "change" and "same" trials (see the upper right-hand panel of Figure 4). Accuracy in "change" trials was constantly lower than that in "same" trials across array sizes. This observation was consistent with the results of a two-way (2 trial types X 4 array sizes) ANOVA, showing a significant main effect of trial type ($F(1, 14)=28.12$, $p <.001$, $\eta_p^2 = .67$), a significant main effect of array size ($F(3, 42)=19.70$, $p<.001$, $\eta_p^2 = .58$), but no interaction between array size and trial types.

Correlation between K_2 and Cowan's K in Experiment 1a and 1b. To supplement the group mean analysis, the correlation between K_2 and Cowan's *K* across array sizes in Experiment 1a and 1b was also investigated. An average capacity estimate, on the basis of the non-rational strategy and the all-or-none assumptions was first calculated across the multiple-choice identity recall and location recall for each participant. A Pearson's correlation coefficient was then computed between the average K_2 across both multiple-choice recall tasks and Cowan's K from the change detection task for Experiment 1a and 1b. The results were plotted in Figure 5, r = .73 in Experiment 1a with English consonants as stimuli and r = .77 in Experiment 1b with Chinese characters as stimuli. The fairly high correlation between K_2 and Cowan's K in both Experiment 1a and 1b suggests that the validity of analyses based on group means should be able to extend to the individual level.

Response Time Analysis across Experiment 1a and 1b. In order to check whether participants systematically changed their strategy across Experiment 1a and 1b and across response modes, the response time data were analyzed. A two-way (2 Experiments X 3 Response modes) between-participants ANOVA on mean response times showed neither main effect ($F(1,28) = .081$, $p = .78$, $\eta_p^2 = .002$) of experiment nor interaction ($F(2, 56) = 2.91$, $p = .06$, $\eta_p^2 = .09$) between experiment and response mode factors, suggesting that Experiment 1b had not encouraged a strategy different from that in Experiment 1a. There was indeed a significant main effect of response mode $F(2, 56)=57.30$, $p <.001$, $\eta_p^2 = .67$ and a post hoc Newman-Keuls test attributed the cause of such response mode main effect to a longer response time in the identity recall mode than in two other response modes (p's < .001), suggesting that searching a set of identity options for the critical item did take longer than making a "change" or "same" judgment on a probe item or indicating the critical item's location.

Modeling data from Experiment 1a and 1b. A tentative multinomial model was developed to account for data from both Experiment 1a and 1b. This model released the all-or-none assumption to accommodate partially encoded VWM representations in Experiment 1b. The model can be illustrated with a tree structure shown in Figure 6 & 7. Assumptions of this model can be delineated as follows. *K* items, regardless of their resolution states, can be encoded into VWM. If a critical item is in VWM, its identity representation can be in either high or low resolution. A critical item in high resolution always leads to a correct response, whereas a critical item in low resolution may lead to a correct response with a probability contingent on the task type. Specifically, the "same" trials have a probe item identical to the critical item, providing the strongest retrieval support, whereas the "change" trials have a probe different from the critical item, providing the weakest retrieval support. The retrieval support for a multiple-choice recall trial should be intermediate, as the critical item is available in the option set but the non-critical items may also distract attention or cause interference. Theoretically, the probability of a correct response based on partial information may be relatively lower for the "change" trials and the multiple-choice recall trials but it should be higher than the chance level. When a critical item is not in VWM, a correct response is possible only by guessing.

Let *K*, *Y* and *N* denote, respectively, the entire VWM capacity, the number of low-resolution items (with the restrictions $Y \leq K$ and $Y \leq N$), and the array size. These are probability parameters for a correct response to low-resolution representations: 1) X_1 is the probability of responding "change" in a "change" trial; 2) X_2 is the probability of responding "same" in a "same" trial; and 3) X_3 is the probability of a correct response in a multiple-choice recall task (with the restrictions $(X_1 + X_2)>1$ and $X_3 >1/N$). In addition, let g and $(1-g)$ denote the probability of guessing "change" and "same" in a change detection task, respectively.

With denotations specified above, the hit rate and correct rejection rate in the change detection task can be formulated in equations (12) ~ (16) as below.

When $K \leq N$

$$ht=(K-Y)/N+(Y/N)*X_1+(1-K/N)*g \qquad (12)$$

$$cr=(K-Y)/N+(Y/N)*X_2+(1-K/N)*(1-g) \qquad (13)$$

When $N<K<N+Y$

$$ht=(K-Y)/N+(Y/N)*X_1 \qquad (14)$$

$$cr=(K-Y)/N+(Y/N)*X_2 \qquad (15)$$

When $K \geq N+Y$

$$ht=cr=1 \qquad (16)$$

The rationale of the hit rate formulation can be described as follows. When $K \leq N$, given that there are Y items of low resolution in the sample array of N items and a critical item is randomly selected, the probability of the critical item being in low resolution is Y/N. Consequently, given a VWM capacity of K and a sample array of N items, $(K-Y)$ items would be in high resolution and the probability of the critical item being in high resolution is $(K-Y)/N$. As a result, the probability of a critical item with no feature in VWM would be $(1-Y/N-(K-Y)/N)=(1-K/N)$. Since a high-resolution critical item presumably leads to correct response, the probability of having a correct response from a high-resolution critical item is $(K-Y)/N$. Likewise, with X_1 being the probability of responding “change” to a low-resolution critical item in a “change” trial, the probability of having a correct response to a low-resolution critical item is $(Y/N)^* X_1$. Finally, the probability of responding “change” by chance to a critical item that has no feature in VWM is $(1-K/N)^*g$. A similar logic can be applied to the correct rejection rate formulation, except that X_1 and g in equation (12) are replaced by X_2 and $(1-g)$, respectively.

When $K-Y<N$ (i.e., $K<N+Y$), performance will depend on Y as in Equations (14) and (15). If, however, $K \geq N+Y$, then all items will be fully encoded and performance will be perfect.

Similarly, the accuracy in the multiple-choice recall task can be specified in equation (17)~(19).

When $K \leq N$,

$$Ac=(K-Y)/N+(Y/N)^*X_3+(1-K/N)^*(1/N) \qquad (17)$$

When $N<K<N+Y$

$$Ac=(K-Y)/N+(Y/N)^*X_3 \qquad (18)$$

When $K \geq (N+Y)$,

$$Ac=1+(Y/N)^*X_3 \qquad (19)$$

The rationale behind the formulation is similar to those of the hit rate and correct rejection rate. Note that the guessing rate is $1/N$, meaning one out of N options, rather than g or $(1-g)$.

It can be shown that these equations provide a fairly good fit to the data. Preliminary model fitting was carried out by adjusting free parameters in the equations (12), (13) and (17) by eye to match the data. For the multiple-choice recall task, only the identity recall data were used, as they were very similar to those of the location recall. The parameter values for both Experiment 1a and 1b are reported in Table 3 and the observed data and predicted values can be found in Table 4 and were also plotted in Figure 8. It is especially noteworthy that the fits were obtained using the same parameter values for both Experiments 1a and 1b, except for the parameter *Y*, the number of items that are in VWM but in low resolution. That parameter value was 0 in Experiment 1a and 3 in Experiment 1b.

As the partial information model was fitted at the group mean level in the above analysis, the fitting at an individual level was also investigated through plotting individual residuals from the predicted group mean (See Figure 9). The residuals were mostly symmetrically distributed around 0, suggesting the fitting at individual level was consistent with that at the group mean level.

Another model based on a partial information account but a "rational" response strategy was also tried out. Such a model only needs to modify equation (17) into $Ac=(K-Y)/N+(Y/N)*X_3+(1-K/N)*(1/(N-K))$. A trial and error fitting process was unable to simulate even the pattern of the observed data.

Auxiliary experiment. An additional experiment was administered but not reported in the present paper because the data pattern was disorderly. This experiment was similar to Experiment 1b, but the change detection task included a type of "change" trials, known as "binding-change" or "feature-switch" trials, in which the probe item was one of the stimuli from a different location in the sample array rather than an outside item. A derivative VWM capacity estimate based on the “binding-change” trials is logically complicated (Alvarez & Thompson, 2009). It was found that the binding-change trials showed higher accuracy performance than other VWM trials, suggesting participants might be able to take advantage of some information from non-critical items in VWM under some circumstances. In particular, the participant might know that a probe has changed from the studied array because he or she remembered that the probe item appeared at a different location in the array.

Discussion

Experiment 1a and 1b enriched our knowledge about the boundary conditions for the all-or-none assumption in visual working memory research. As Experiment 1a showed, the all-or-none assumption is not limited to the change detection paradigm, but equivalently viable for multiple-choice identity or location recall procedure as well. Data from different procedures but with the same all-or-none assumption showed that VWM capacity estimates ascended to asymptote at the level of about 4 items. This asymptote was not caused by any ceiling effect, as it had room to reach higher given the largest array size was 6. Therefore, Cowan’s K was validated by converging evidence from two other derivative VWM capacity estimates that all shared the all-or-none assumption.

It is noteworthy, though, that the validation of Cowan’s K required a departure from the rational strategy in which the participant uses all of the information that he or she knows. That assumption led to capacity estimates that were very divergent for the different response modes (See Figure 2). In contrast, the assumption that the participant does not use information in VWM about non-critical items led to nicely convergent functions (Figure 3, left-hand panel).

The non-rational use of visual working memory probably echoes the somewhat-efficient conditioning in perception of letters and words (Rouder, 2004). In an investigation of the mechanism that rapidly presented letters and words are better identified when there are fewer available choices, Rouder (2004) found that observers are "neither as efficient as they would be by using ideal conditioning nor as inefficient as they would be by using choice-set restrictions when guessing." To some extent, using ideal conditioning through normalizing activation by the total activation of available choices (e.g. Keren & Baggen, 1981; Oden & Massaro, 1978) is comparable to the rational use of visual working memory in the present project, whereas guessing among all available options (e.g. McClelland & Rumelhart, 1981) can be likened to the non-rational use of visual working memory. However, for an automatic perception process, a non-rational processing is confounded with the partial representation of stimuli as the cause of inefficient conditioning. In contrast, the present project successfully separated and demonstrated these two mechanisms in visual working memory tasks.

Despite no restriction by test procedures, the nature of stimuli does impose a boundary condition for the all-or-none assumption. Experiment 1b showed that different procedures with the same all-or-none assumption showed discrepancies in VWM capacity estimates, when Chinese characters replaced English consonants as stimuli for participants who did not speak Chinese. With the present methods, it is reasonable to assume that both English consonants and Chinese characters in a sample array should have been encoded in two dimensions: the identity features and the location features. The critical difference between these two types of stimuli may be in how the identity information was encoded. Given that English consonants were very familiar to the participant, identity features of English consonants could be registered into working memory with either a single conceptual feature or a single phonological feature. In contract, a 2-or-3-stroke Chinese character might require at least two orthographic features at the same time to sufficiently register its identity. Given that features of the same dimension may compete for capacity and attention resources in encoding (Wheeler & Treisman, 2002), Chinese characters may not be encoded in an all-or-none manner; in other words, at least some Chinese characters in the sample array might be partially represented in VWM. The multiple-choice recall procedure always provided the intact identity or location information of the critical item so that it only had to be recognized, whereas the change detection procedure gave that kind of information only for the "same" trials. Consequently, partially-encoded critical items might lead to correct responses more often in the multiple-choice procedures. Indeed, this description of processing matches our multinomial model and explains 1) the discrepancy between the recall and detection procedures in accuracy, and 2) the bias favoring "same" trials over "change" trials in the change detection procedure.

The predicted values of mean proportion correct by the model apparently fitted the observed data fairly well. The strikingly different patterns between

Experiment 1a and 1b were satisfactorily reconciled by changing only a single parameter *Y*, the number of low-resolution items in VWM (See Table 3). Thus, it is reasonable to hypothesize that the difference between data patterns in Experiment 1a and Experiment 1b was solely caused by the number of low-resolution items in VWM, with Experiment 1a having no low-resolution item at all but Experiment 1b having about 3 per trial. More importantly, this model showed that a constant VWM capacity is possible regardless of visual complexity in stimuli, with a VWM capacity of 3.8 items across participants in both experiments. The relationship between the predicted values X_1, X_2, and X_3, with $X_2 > X_3 > X_1$, is reasonable because a probe item in a “same” trial with the change detection response mode provided the most retrieval support to a low-resolution critical item, whereas a probe item in a “change” trial did the least. This was also consistent with the fact that guessing biased toward responding “same” (i.e., g = .3) in a change detection task.

One challenge to the discrepancy in VWM capacity estimates between different procedures across Experiment 1a and 1b is that participants might choose a “non-rational” strategy (i.e., not making full use of available information in VWM) in Experiment 1a, but switch to a “rational” strategy (i.e., making full use of available information in VWM) in Experiment 1b. In other words, for experiment 1a, the VWM capacity based on the multiple-choice recall procedures should be estimated with the K_2 formula, whereas for experiment 1b, the K_1 formula should be used instead. Therefore, there will be little discrepancy between Cowan’s K and two derivative capacity estimates any more. However, this argument is not reasonable and also lacks empirical support. For one thing, it is more demanding to apply the “rational” in Experiment 1b than the “non-rational” strategy in Experiment 1a, as it is obviously harder to hold K Chinese characters than English consonants in mind for a while until they can be eliminated from the option set. For another, this argument lacks support from the response time data. Theoretically, it involved more steps of processing to carry out the “rational” strategy when participants made a guess by first eliminating multiple non-critical items from the option set and then randomly selecting one among the rest of the option set. The response time associating with the “rational” strategy is expected to be longer in Experiment 1b. In other words, the multiple-choice response modes should take longer in Experiment 1b than in Experiment 1a. In fact, the non-significant main effect of experiment and interaction between experiment and response mode indicated that it was unlikely that participants systematically changed their response strategies across Experiment 1a and 1b. If participants in Experiment 1b did rely on the “rational” strategy, a main effect of experiment factor should be observed. Moreover, the significant main effect of response modes confirmed my assumption of the response time analysis that mean response time was sensitive enough to distinguish different response strategies. Therefore, the response time analysis should help to rule out the possibility that switching between “rational” and “non-rational” strategies can account for the discrepancy among VWM capacity estimates across Experiment 1a and Experiment 1b.

In addition, accuracy data from Experiment 1a and 1b suggested that identity and location features were integrated fairly well regardless of complexity of the visual stimuli. In both experiments, participants had comparable accuracy with identity and location recall trials, although they did much better with English consonants than with Chinese characters. This pattern implies that identity and location features can be bound even though each feature dimension might not have been fully encoded and the binding between identity and location supports bidirectional retrieval, that is, the identity feature can be used to cue the retrieval of the location feature and the reverse is equally efficient.

Although our theoretical model provides an account of the data, there are still some important unknowns about the difference between VWM for English consonants versus unfamiliar Chinese characters. The remaining two experiments address two of these unknowns. Experiment 2 explored the extent to which the encoding of English consonants depended on phonological encoding, and Experiment 3 explored the extent to which the poor memory for Chinese characters could be ameliorated using a longer encoding period.

Experiment 2: VWM Tasks with 3 Response Modes and English Consonants as Stimuli, Under Articulatory Suppression

Experiment 1a and 1b showed that the all-or-none assumption yielded consistent VWM capacity estimate for different procedures when English consonants were used as stimuli, but that was no longer the case when the stimuli were Chinese characters. One hypothetical account for this phenomenon is that the English consonants are better known but another, slightly different account is that the ability to use phonological encoding (i.e., to pronounce the letter silently) is important for working memory. In Experiment 2 an articulatory suppression procedure from Chen & Cowan (2009) was adopted as a secondary task concurrent to the entire course of the sample array presentation with English consonants. This articulatory suppression procedure was expected to interfere with the phonological encoding or rehearsal (Baddeley, 1986; Baddeley & Hitch, 1974), so that the familiarity of English consonants could still play a role, but not through phonological means.

Method

Participants, Stimuli, Conditions & Procedure. Experiment 2 had 30 participants (23 females and 7 males). Methods regarding stimuli and conditions were identical to those in Experiment 1a. In brief, it used English consonants as stimuli and had 3 VWM tasks with different response modes.

The procedure of Experiment 2 was highly similar to that of Experiment 1a. The only difference between the two experiments was Experiment 2's inclusion of an articulatory suppression procedure that concurred with the fixation signal display, the sample array display and the clear screen delay (See Figure 10).

Upon receiving the fixation signal display after pressing the SPACE BAR in each trial, participants heard from two loudspeakers flanking the computer screen a voice recording of the word “the” twice per second continuously. Meanwhile they were supposed to say aloud the same speech at about the same pace over a period of about 2000 msec, presumably while encoding the sample array. The voice playback stopped and participants ceased to articulate as soon as the mask display occurred. Participants were allowed to drink water during breaks between trials if they felt thirsty from articulation.

Results

In Experiment 2, each of 30 participants performed 480 VWM trials under articulatory suppression concurrent with the sample array display and responded in three different modes. Descriptive statistics on VWM capacity estimates and accuracies of various conditions are reported in Table 5. The mean VWM capacity estimates ranged from 2.67 to 3.41 across conditions.

As Figure 11 shows, three VWM capacity estimates, including Cowan’s K and the other two derivative K_2s, started at about 2.7 for an array size of 3 and rose a little with increased array sizes, but generally leveled off after the array size exceeded 4, as there were no significant difference between array sizes 5 and 6. A two-way (3 response modes x 4 array sizes) ANOVA on VWM capacity estimates yielded a significant main effect of response mode ($F(2, 58)=5.57$, $p<.01$, η_p^2 =

.16), a significant main effect of array size ($F(3, 87)=12.65$, $p <.001$, $\eta_p^2 = .30$), but no significant interaction between response mode and array size.

Nevertheless, the capacities observed in Experiment 2 suggest a two-stage relationship among three Ks. Post hoc Newman-Keuls tests showed that when array sizes were not greater than 4, there were no significant differences among the three Ks. When array sizes were greater than 4, there was still no difference between two derivative K_2s, but Cowan's K (for the change-detection task) dropped below both two derivative K_2s (for the identity and location tasks), with the difference reaching statistical significance at array size 5 (p's < .05) though not quite significant at array size 6. The statistical power appears to have been insufficient to observe what appears to be an interaction between response mode and array size.

Discussion

Experiment 2 was intended to test the partial information account that was proposed to reconcile the different data patterns observed from Experiment 1a and 1b. To this end, Experiment 2 utilized an articulatory suppression procedure to degrade the phonological codes of English consonants' identity feature. It was expected that, although with the same stimuli (i.e., English consonants) of Experiment 1a, the data of Experiment 2 should resemble that of Experiment 1b with Chinese characters as stimuli. The result turned out to be an intermediate pattern between Experiment 1a and 1b, only somewhat consistent with the prediction.

Three VWM capacity estimates showed discrepancy after the array size exceeded 4 items. It is possible that the phonological loop was called into service only after the focus of attention had been filled up (Cowan, 2001). Within the capacity of the focus of attention, the identity feature of English consonants could be encoded in conceptual codes, so that articulatory suppression didn't succeed in preventing the all-or-none coding of stimuli. Therefore, the three Ks gave consistent measures of the VWM capacity. Once the array size went beyond about 4 items and phonological codes had to be used to encode the categories of the stimuli, articulatory suppression then started to work against the all-or-none assumption. As a result, the three VWM capacity estimates then led to divergent measures.

Alternatively, articulatory suppression might impose a dual-task effect on visual working memory. Studies on dual task performance have suggested that a secondary task could be detrimental to a primary memory task if the two tasks concur during the encoding phase of memory (e.g., Naveh-Benjamin, Craik, Guez, & Dori, 1998; Naveh-Benjamin, Guez, & Marom, 2003). An attention-based interference between visual working memory and articulatory suppression might become severe when the array size exceeds 4 items, that is, the capacity of the focus of attention (Cowan, 2001), so that the encoding resolution of English consonants was significantly degraded and resulted in a divergence between different capacity estimates at array size of 5.

Experiment 3: VWM Tasks with 3 Response Modes, Chinese Characters as Stimuli, and Two Display Durations in the Sample Array

The divergent results on VWM capacity estimates between Experiment 1a and 1b indicate a boundary condition for the all-or-none assumption. One explanatory hypothesis attributes the failure of the all-or-none assumption to partial information that the representation of complex objects such as Chinese characters contains in working memory. Experiment 2 tested this partial information account through making Experiment 1a more like Experiment 1b. Conversely, in Experiment 3, the partial information hypothesis was examined through making Experiment 1b more like Experiment 1a. Experiment 1b was modified to allow more time for encoding the sample array. It was expected that with more information from the sample array encoded into VWM, the validity of the all-or-none assumption could be somewhat restored even with Chinese characters as stimuli.

Method

Participants, stimuli, & procedure. Experiment 3 had 24 participants (13 females and 11 males). The method of Experiment 3 was almost identical to that of Experiment 1b, except that Experiment 3 had one half of the trials with a display duration of 500 msec and the other half with 1000 msec for the sample array regardless of response modes. In brief, Experiment 3 used Chinese characters as stimuli and had 3 VWM tasks with different response modes, and a total of 480 trials, half of which allowed 500 msec to see the sample array and the other half of which, 1000 msec.

Results

Descriptive statistics on VWM capacity estimates and accuracies for two display durations in Experiment 3 can be found in Table 6. The mean VWM capacity estimates ranged from 1.35 to 2.34 across all conditions. Because the results were somewhat noisy, I collapsed the data across set sizes in order to get a simpler picture of the effect of display duration on capacity estimates.

The estimates are shown in Figure 12 and were analyzed in a two-way (2 display durations X 3 response modes) ANOVA. There was a significant main effect of display duration ($F(1, 23)=16.46$, $p <.001$, $\eta_p^2 = .42$), and a significant main effect of response mode ($F(2, 46)=5.40$, $p <.01$, $\eta_p^2 = .19$) but no significant interaction between response mode and display duration. Nevertheless, the figure suggests that there is a tendency for the capacity estimates to be more similar across different response modes with the longer display duration.

Separate one-way analyses for the two display durations confirmed this suggestion. For the 500-msec display duration there was an effect of the response mode, $F(2, 46)=4.97$, $p<.05$, $\eta_p^2 = .18$; post-hoc Newman-Keuls tests indicated that Cowan's K in the change detection mode was significantly lower than both derivative K_2s in the multiple-choice recall mode, p's<.05. In contrast, for trials with a 1000-msec display duration, there was no significant difference among three VWM capacity estimates, $F(2,46) = 1.32$, $p= 0.28$, $\eta_p^2 = .05$.

Discussion

In order to examine the partial information account, the purpose of Experiment 3 was to facilitate the encoding process of Chinese characters by giving an extra 500 msec when they were presented on the sample array and investigate if this manipulation helped to restore the validity of the all-or-none assumption. It was expected that, even though Chinese characters were used as stimuli, the quantitative relationship among different VWM capacity estimates should resemble that in Experiment 1a, in which English consonants were used as stimuli. Although the pattern of results was noisy, there was a trend in the appropriate direction. Specifically, a significant difference between the mean capacity estimates for different response modes existed with a 500-msec display duration, but not with a 1000-msec display duration. In summary, the validity of the all-or-none assumption for unfamiliar Chinese characters appears to have been improved, though not fully restored, with a longer display duration in this experiment.

General Discussion

Summary

The present project started by exploring the boundary conditions for the all-or-none assumption that underlies a popular VWM capacity estimate, Cowan's K. Such exploration was carried out along two directions. Experiment 1a addressed whether the all-or-none assumption can be extended beyond the conventional change detection paradigm, and Experiment 1b examined whether the all-or-none assumption can be extended to procedures with complex visual objects.

The findings from Experiment 1a and 1b suggested that 1) the all-or-none assumption enjoyed converging and supportive evidence from different VWM test procedures, as long as the visual stimuli were simple and familiar, but 2) when complex and novel stimuli were used, the all-or-none assumption led to significant discrepancies across different test procedures.

Mathematical modeling led to two additional, important conclusions. First, the all-or-none model only converged across response modes when a non-rational model was used. In a rational model, a participant in the identity or location recall task who does not remember the probed item still can make use of memory of other items to narrow the choices. Instead, the results suggested that a non-rational method was used in which the participant either knew the probed item or guessed randomly, not using knowledge of other items to narrow the choices. Perhaps this occurs because the use of information from multiple items at once requires switching attention from one item to another (Oberauer, 2002), which can cause interference. The use of a non-rational strategy is a well-known phenomenon, e.g., in the situation in which participants match the probabilities of making different responses to the probability that each of these responses is rewarded, rather than using the optimal strategy of always making the response choice with the highest probability of reward (Estes, 1962).

Second, as shown in Figure 8, it was possible to account for performance on Chinese characters using the non-rational model with the same parameter values as English consonants, except that there was the need for an extra parameter expressing the number of characters for which only partial information was available (set to zero for English consonants, but not for Chinese characters). The success of this modeling endeavor suggests that the proper extension of the model is one in which there is a fixed capacity, but in which the resolution of items filling that capacity can vary. If the resolution is too little for the task then the information cannot be used reliably. This conception is highly consistent with other recent work (Awh et al., 2007).

The project further investigated the failure of all-or-none assumption caused by complex visual objects and a partial-information account for the assumption's breakdown was proposed and tested in Experiments 2 and 3. These studies showed that preventing articulation weakens the fit of the all-or-none assumption for English consonants (Experiment 2), and that allowing a longer display duration for extended encoding somewhat strengthens the fit to the all-or-none assumption for Chinese characters (Experiment 3).

When Does the All-or-none Assumption Work?

The most remarkable finding from the present project was that the all-or-none assumption can be extended to other VWM test procedures than the conventional change detection paradigm and Cowan’s K successfully converged with derivative VWM capacity estimates from other procedures, when simple and familiar objects were used as stimuli.

The issue about how VWM capacity is limited has been intensively debated in recent years. VWM capacity could be limited in terms of a number of discrete slots (Cowan, 2001; Cowan & Rouder, 2009; Luck & Vogel, 1997; Zhang & Luck, 2008) or alternatively, in terms of a pool of dynamic cognitive resources (Bays & Husain, 2008; Davis & Holmes, 2005; Wilken & Ma, 2004). Findings from the present project appeared more in line with the “slots” theory of VWM capacity, as it showed that the all-or-none assumption for working memory representations is practically legitimate under some circumstances.

The successful extension of the all-or-none assumption across different VWM test procedures also carried significant implication for methodology development in cognitive psychology. The present project tested a target hypothesis (i.e., VWM representations is all-or-none) that is almost impossible to be ascertained directly. An indirect approach was adopted by examining the testable inferences (i.e. the convergency of VWM capacity estimates with different response modes) of the intangible target hypothesis. An intuitive analogy of this idea is like the following. You want to test the hypothesis that a farmer uses buckets of the same exact size to carry grapes sometimes and plums sometimes to process dried fruits. You do not see the bucket but you find little piles of dried grapes (raisins) and dried plums (prunes) on the sunning ground. They are in different sizes but you calculate that when the dried fruit was not yet dried, the piles of grapes and piles of plums fit into exactly the same volume. Suppose you also know that equal numbers of buckets were used to transport both grapes and plums. This way, you should have strengthened your constant-bucket-size theory without having observed the buckets. Similarly, if a chemist has to prove both diamond and graphite are solely made up of carbon, probably there is no direct way to see carbon atoms in both materials. However, an indirect and easy way to prove the hypothesis is to simply burn each material in pure oxygen and show the only product from both reactions is carbon dioxide, with identical characteristics regardless of which source.

When Does the All-or-none Assumption Not Work?

With the same inference method, the present project also identified situations in which the all-or-none assumption does not work very well. As Experiment 1b showed, complex visual objects such as Chinese characters led to divergent VWM capacity estimates across different response modes, suggesting that working memory representation of the stimuli might not be all-or-none. The rationale is like this. The multiple-choice recall mode provided retrieval support for every trial by showing all intact stimuli as options in the test array display, whereas the change detection mode gave similar support only for the “same” trials. Such support should not be helpful unless the representation of a critical item was partially available in working memory. In other words, if this support helps, the working memory representation of a critical item should not be all-or-none. In Experiment 1b, the observed divergence among different VWM estimates indicated that the multiple-choice recall modes indeed benefited more from the support than the change detection mode. Therefore, this finding reflected the existence of partial information in working memory representations, suggesting that the all-or-none assumption fell apart with Chinese characters. In addition, the divergence among different VWM capacity estimates was not the only evidence against the all-or-none assumption in Experiment 1b. The declining trend of those VWM capacity estimates with increasing array sizes was not right either, as the trend of VWM estimates in Experiment 1a was ascending toward an asymptote. One plausible explanation for that could be that the VWM capacity was even smaller than the smallest array size (i.e., 3 Chinese characters) in Experiment 1b, so that there was no room for the VWM capacity estimate to ascend with increasing array sizes. However, it is still hard to explain why the VWM capacity estimate was not simply leveling off rather than declining if the all-or-none assumption is really true.

Experiment 2 showed another situation in which the validity of the all-or-none assumption may be at risk. When English consonants were presented concurrent with an articulatory suppression task, different VWM test procedures failed to yield converging estimates at a larger array size, suggesting a possible breakdown of the previously well-integrated representation of the English consonants in visual working memory. This finding indicated that whether a visual stimulus can be encoded in an all-or-none manner is not only depending on its visual complexity, but also on the process of encoding, which to some extent seems consistent with a proposal by Wheeler and Treisman (2002) that features of a visual object requires focused attention to create and maintain the binding over time, as it is vulnerable to interference.

Why Does the All-or-none Assumption Break Down?

A partial information account was proposed to address the reason why the all-or-none assumption fails under some circumstances. This hypothetical account was able to explain different data patterns in Experiment 1a and 1b in a satisfactory manner, because the only factor that changed systematically across Experiment 1a and 1b was the visual complexity of stimuli.

The partial information account was further tested in Experiment 2 and 3 through manipulating factors that affected the encoding processes. When encoding of English consonants in the sample array was interfered by articulatory suppression, the VWM capacity estimate data pattern deviated from one consistent with the all-or-none assumption. In contrast, when encoding of Chinese characters was facilitated with a longer sample array display, the VWM capacity estimate data pattern regressed to one more consistent with the all-or-none assumption.

There could be at least three reasons why complex and novel stimuli such as Chinese characters for speaking-no-Chinese observers failed to form an all-or-none working memory representation. First, it takes more attention resources to fully perceive a complex object as it usually had more features. Second, if a stimulus is novel, its exact representation is not available in long-term memory.

So, a temporary substitute has to be used as its representation in working memory, which could cause considerable confusion down the road of further processing. In particular, it might be necessary to hold in mind several features separately rather than retaining them as a fused whole as in the case of familiar characters or consonants. Third, when a character is unfamiliar it does not lend itself to covert verbal articulation the way that a known consonant does, which eliminates one powerful mode of retaining the information.

Theoretical Implications of the Present Project

The present project draws attention to the importance of examining assumptions in cognitive psychology. Almost all laws of the physical world have their applicable domains or boundary conditions. Once a law goes beyond its domain it will lose its validity, just as laws of classical mechanics fail when they are applied to the movement of atomic particles. In contrast, in the field of psychology, boundary conditions of psychological theories or models are often left unexplored, probably due to either technical difficulties or negligence. Psychological research sees many prolonged but hardly settled debates, such as nature-based versus nurture-based cognitive development, decay-based versus interference-based forgetting, discrete-slots-based versus dynamic-resources-based VWM capacity and so on. Psychologists tend to fight against one theory in favor of the other. However, the truth might well be that conflicting theories could be, at least pragmatically, reconciled by taking the boundary conditions of their assumptions into consideration. With more attention and resources allocated to exploring the boundary conditions of theories, modern psychology as an emerging scientific discipline might be able to make more substantial progress in the future.

This paper reports an interesting finding of human strategy use. Research in cognitive psychology tends to assume that human participants rationally process mental information in an experimental setting. In other words, it is widely believed that cooperative experimental participants would maximize their mental resources and available information to achieve best performance in a cognitive task. However, the present project showed that this might not be the case. In Experiment 1a, participants seemed to utilize a “non-rational” strategy, with which they did not make full use of information held in VWM to achieve better performance. As for its reasons, it might be the mental difficulty in carrying out the “rational” strategy that prevented them from doing so. Alternatively, although the “rational” strategy was achievable, participants might prefer a strategy of less effort so that they could complete the experiment sooner or in a more comfortable manner.

More importantly, this paper proposes the partial information account and supports it with a viable multinomial model. It has been controversial how visual complexity affects VWM performance. A number of hypotheses have been proposed, including 1) the capacity-reduced view by Alvarez & Cavangh (2004), suggesting that the number of objects held in visual working memory is an inverse function of the information load per object but ranges from 0 to 4 or 5 in discrete units; 2) the encoding-bottleneck view by Eng et al.(2005), propounding that object complexity may impose an encoding bottleneck, with which fewer complex objects can be encoded into visual working memory; and 3) the resolution-reduced view by Awh et al (2007), arguing that the reduced visual change detection performance with complex objects is caused by the reduced resolution rather than the reduced capacity in visual working memory. Meanwhile, the popular VWM capacity index, Cowan's K(Cowan, 2001), has yet to provide a solution to the issue. To address that, the present project first identified a common limitation in these studies that all of them relied on the change detection paradigm to estimate VWM capacity. This procedure per se does not reveal the failure of the all-or-none assumption and the partial representation of complex objects in VWM. Based on that finding, this paper proposes a set of equations (12)~(16) to supplement the Cowan's K formulation (equations (1)~(4)). Such revision not only extends the validity of Cowan's K to complex stimuli, but also redefines the notion of visual resolution, which is based on number of features in VWM rather than in terms of perceptual similarity (Awh et al., 2007) or preciseness (Zhang & Luck, 2008). In addition, the multinomial model successfully depicts a quantitatively specific mechanism of how visual complexity influences VWM performance, provided a constant and discrete VWM capacity.

Nevertheless, there are still issues that the present project did not fully address. For example, when the all-or-none assumption cannot be met and Cowan's K diverges from other derivative VWM capacity estimates, which VWM capacity estimate, K_2 based on the multiple-choice recall procedure or Cowan's K based on the change detection procedure, would be a better index of

VWM capacity? Also, although a systematic adoption of the “rational” strategy with the multiple-choice recall procedure was not observed in Experiment 1a & 1b, could it still be possible that discrepancy between Cowan’s K and two derivative Ks was primarily due to occasional application of the “rational” strategy by some participants? Will it provide us with deeper insights into participants’ response strategies through manipulating the response option set in the multiple-choice recall task? To answer these specific questions calls for future efforts and will definitely advance the research on human visual working memory capacity.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106-111.

Alvarez, G. A., & Thompson, T. W. (2009). Overwriting and rebinding: Why feature-switch detection tasks underestimate the binding capacity of visual working memory. *Visual Cognition., 17*(1-2), 141-159.

Atkinson, R., & Shiffrin, R. (1968). Human memory: A proposed system and its control processes, *Spence, Kenneth W; Spence, Janet T (1968) The psychology of learning and motivation: II xi, 249 pp Oxford, England: Academic Press*.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622-628.

Baddeley, A. (1986). *Working memory*. New York, NY: Clarendon Press/Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science, 321*(5890), 851-854.

Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon Press.

Broadbent, D. E. (1975). Cognitive psychology and education. *British Journal of Educational Psychology, 45*(2), 162-176.

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: The limit in words retained without covert articulation. *The Quarterly Journal of Experimental Psychology, 62*(7), 1420 - 1429.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement, 33*, 107-112.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552.

Cowan, N. (1995). *Attention and memory: An integrated framework. Oxford psychology series, No. 26*. Attention and memory: An integrated framework xv, 321 pp New York, NY, US: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-185.

Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK Psychology Press.

Cowan, N., & Rouder, J. N. (2009). Comment on "Dynamic Shifts of Limited Working Memory Resources in Human Vision". *Science, 323*(5916), 877c.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*(4), 450-466.

Davis, G., & Holmes, A. (2005). The capacity of visual short-term memory is not a fixed number of objects. *Memory & Cognition, 33*(2), 185-195.

Eng, H. Y., Chen, D. Y., & Jiang, Y. H. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12*(6), 1127-1133.

Estes, W. K. (1962). Learning theory. *Annual Review of Psychology, 13*, 107-144.

Hannon, B., & Daneman, M. (2006). What Do Tests of Reading Comprehension Ability Such as the VSAT Really Measure?: A Componential Analysis. In A. V. Mittel (Ed.), *Focus on educational psychology* (pp. 105-146). Hauppauge, NY: Nova Science Publishers.

James, W. (1890). *The principles of psychology*. NY: Henry Holt and Company.

Kawasaki, M., Watanabe, M., Okuda, J., Sakagami, M., & Aihara, K. (2008). Human posterior parietal cortex maintains color, shape and motion in visual short-term memory. *Brain Research, 1213*, 91-97.

Keren, G., & Baggen, S. (1981). Recognition models of alphanumeric characters. *Percept Psychophys, 29*(3), 234-246.

Luck, S. J. (2007). Visual short term memory. *Scholarpedia, 2*(6), 3328.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review, 88*(375-407).

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77-94.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Henry Holt and Co.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Models of working memory: Mechanisms of active maintenance and executive control xx, 506 pp New York, NY, US: Cambridge University Press.

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11*(2), 296-301.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology-Learning Memory and Cognition, 24*(5), 1091-1104.

Naveh-Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*(7), 1021-1035.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 411-421.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review, 85*(3), 172-191.

Olson, I. R., Moore, K. S., & Drowos, D. B. (2008). The contents of visual memory are only partly under volitional control. *Memory & Cognition, 36*(7), 1360-1369.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*(10-12), 1227-1268.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*(4), 369-378.

Phillips, W. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283-290.

Rouder, J. N. (2004). Modeling the effects of choice-set size on the processing of letters and words. *Psychological Review, 111*(1), 80-93.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975-5979.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225-237.

Sakai, K., & Inui, T. (2002). A feature-segmentation model of short-term visual memory. *Perception, 31*(5), 579-590.

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology-General*, *136*(4), 663-684.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide.* . Pittsburgh: Psychology Software Tools Inc.

Song, J. H., & Jiang, Y. H. (2006). Visual working memory for simple and complex features: An fMRI study. *Neuroimage, 30*(3), 963-972.

StatSoft, I. (2008). STATISTICA (data analysis software system) (Version 8.0).

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751-754.

Vogel, E. K., & Machlzawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748-751.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology-Human Perception and Performance*, *27*(1), 92-114.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology-General, 131*(1), 48-64.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), 1120-1135.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*.

Table 1.

VWM Capacity Estimates, across Three Response Modes and Two Types of Stimuli

Experiment 1a, with English Consonants as Stimuli	**Array Size**			
	3	4	5	6
Change-or-same Detection Mode	2.94(0.08)	3.72(0.31)	3.92(0.79)	4.06(1.08)
Multiple-choice Identity Recall Mode	2.89(0.15)	3.5(0.49)	3.92(0.76)	4.1(1.18)
Multiple-choice Location Recall Mode	2.92(0.15)	3.59(0.48)	3.97(0.82)	4.16(1.22)
Experiment 1b, with Chinese Characters as Stimuli				
Change-or-same Detection Mode	1.66(0.56)	1.57(0.42)	1.32(0.8)	1.18(1.31)
Multiple-choice Identity Recall Mode	2.09(0.57)	1.8(0.64)	1.64(0.54)	1.38(1.01)
Multiple-choice Location Recall Mode	2.19(0.36)	1.8(0.67)	1.69(0.64)	1.63(0.64)

Note. Mean VWM capacity estimates (i.e., Cowan's K and K_2s) (with standard deviations). N = 15 for each experiment.

Table 2.

Proportion Correct in VWM Tasks, Three Response Modes and Two Types of Stimuli

Experiment 1a, English Consonants as Stimuli	**Array Size**			
	3	4	5	6
Change-or-same Detection Mode, Change Trials	0.99(0.02)	0.96(0.06)	0.86(0.12)	0.78(0.18)
Change-or-same Detection Mode, Same Trials	0.99(0.02)	0.97(0.03)	0.92(0.06)	0.9(0.08)
Multiple-choice Identity Recall Mode	0.98(0.03)	0.91(0.09)	0.83(0.12)	0.74(0.16)
Multiple-choice Location Recall Mode	0.98(0.03)	0.92(0.09)	0.84(0.13)	0.75(0.17)
Experiment 1b, Chinese Characters as Stimuli				
Change-or-same Detection Mode, Change Trials	0.67(0.18)	0.62(0.09)	0.49(0.11)	0.48(0.16)
Change-or-same Detection Mode, Same Trials	0.89(0.07)	0.77(0.09)	0.77(0.14)	0.72(0.18)
Multiple-choice Identity Recall Mode	0.8(0.13)	0.59(0.12)	0.46(0.09)	0.36(0.14)
Multiple-choice Location Recall Mode	0.82(0.08)	0.59(0.13)	0.47(0.1)	0.39(0.09)

Note. Mean proportion correct (with standard deviations). N = 15 for each experiment.

Table 3

Experiment	Apparently Best Fitting Parameters					
	K	Y	X_1	X_2	X_3	g
1a	3.8	0	0.45	0.7	0.48	0.3
1b	3.8	3	0.45	0.7	0.48	0.3

Note.

1. K = the number of slots of the entire VWM capacity
2. Y = the number of low-resolution items
3. X_1 = the probability of responding “change” to a low-resolution critical item in a “change” trial
4. X_2 = the probability of responding “same” to a low-resolution critical item in a “same” trial
5. X_3 = the probability of a correct response to a low-resolution critical item in a multiple-choice recall task
6. g = the probability of guessing “change” in a change detection task

Table 4

Simulation of Performance in Mean Proportion Correct

		Array Size			
Experiment 1a, English Letters as Stimuli		3	4	5	6
Change Detection, Hit Rate	Observed Data	0.99	0.96	0.86	0.78
	Predicted Values	1.00	0.97	0.83	0.74
Change Detection, Correct Rejection Rate	Observed Data	0.99	0.97	0.92	0.90
	Predicted Values	1.00	0.99	0.93	0.89
Identity Recall, Accuracy	Observed Data	0.98	0.91	0.83	0.74
	Predicted Values	1.00	0.96	0.81	0.69
Experiment 1b, Chinese Characters as Stimuli					
Change Detection, Hit Rate	Observed Data	0.67	0.62	0.49	0.48
	Predicted Values	0.72	0.55	0.50	0.47
Change Detection, Correct Rejection Rate	Observed Data	0.89	0.77	0.77	0.72
	Predicted Values	0.97	0.76	0.75	0.74
Identity Recall, Accuracy	Observed Data	0.80	0.59	0.46	0.36
	Predicted Values	0.75	0.57	0.50	0.43

Table 5

Experiment 2, Three Response Modes, English Consonants as Stimuli, under Articulatory Suppression

VWM capacity estimates (Ks)	**Array Size**			
	3	4	5	6
Change-or-same Detection Mode	2.72(0.39)	3.13(0.76)	2.98(0.97)	3.01(0.86)
Multiple-choice Identity Recall Mode	2.67(0.38)	3.12(0.6)	3.39(0.81)	3.22(0.86)
Multiple-choice Location Recall Mode	2.78(0.29)	3.26(0.64)	3.41(0.97)	3.21(0.95)
Proportion Correct				
Change-or-same Detection Mode, Change Trials	0.94(0.11)	0.89(0.12)	0.77(0.17)	0.72(0.13)
Change-or-same Detection Mode, Same Trials	0.97(0.05)	0.89(0.09)	0.82(0.1)	0.78(0.12)
Multiple-choice Identity Recall Mode	0.93(0.08)	0.84(0.11)	0.74(0.13)	0.61(0.12)
Multiple-choice Location Recall Mode	0.95(0.07)	0.86(0.12)	0.75(0.15)	0.61(0.13)

Note. Mean VWM capacity estimates (i.e., Cowan's K and K_2s) and mean proportion correct (with standard deviations). N = 30

Table 6

Experiment 3, Three Response Modes, Chinese characters as Stimuli, with Two Sample Array Display Durations

500 msec Display Duration, VWM Capacity Estimates	**Array Size**			
	3	4	5	6
Change-or-same Detection Mode	1.35(0.64)	1.5(1.02)	1.4(1.14)	1.58(1.54)
Multiple-choice Identity Recall Mode	2.09(0.56)	1.8(0.83)	1.37(0.91)	1.55(1.06)
Multiple-choice Location Recall Mode	1.99(0.74)	1.78(0.79)	1.76(0.97)	1.73(0.97)
500 msec Display Duration, Proportion Correct				
Change-or-same Detection Mode, Change Trials	0.69(0.16)	0.58(0.19)	0.54(0.21)	0.55(0.21)
Change-or-same Detection Mode, Same Trials	0.76(0.15)	0.79(0.14)	0.74(0.21)	0.72(0.19)
Multiple-choice Identity Recall Mode	0.8(0.12)	0.59(0.15)	0.42(0.15)	0.38(0.15)
Multiple-choice Location Recall Mode	0.78(0.16)	0.58(0.15)	0.48(0.16)	0.41(0.13)
1000 msec Display Duration, VWM Capacity Estimates				
Change-or-same Detection Mode	1.79(0.54)	1.73(0.86)	1.81(1.14)	2.08(1.66)
Multiple-choice Identity Recall Mode	2.08(0.62)	2.32(0.81)	1.67(1.12)	1.79(1.22)
Multiple-choice Location Recall Mode	2.34(0.46)	2.18(0.87)	1.76(0.97)	1.8(1.09)
1000 msec Display Duration, Proportion Correct				
Change-or-same Detection Mode, Change Trials	0.73(0.14)	0.65(0.16)	0.56(0.17)	0.6(0.23)
Change-or-same Detection Mode, Same Trials	0.87(0.11)	0.79(0.14)	0.8(0.14)	0.75(0.19)
Multiple-choice Identity Recall Mode	0.8(0.14)	0.69(0.15)	0.47(0.18)	0.41(0.17)
Multiple-choice Location Recall Mode	0.85(0.1)	0.66(0.16)	0.48(0.15)	0.42(0.15)

Note. Mean VWM capacity estimates (i.e., Cowan's K and K_2s) and mean proportion correct (with standard deviations). N = 24

Figure Captions

Figure 1. Schematic illustration of procedure in Experiment 1a. Experiment 1b shared the same procedure but with Chinese characters as stimuli.

Figure 2. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_1 (for identity and location recall) as a function of response modes in Experiment 1a with English consonants as stimuli. This analysis was based on the all-or-none assumption and the "rational" strategy model. The error bars indicate standard errors of the mean.

Figure 3. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) as a function of response modes. Left, Experiment 1a, English consonants as stimuli; right, Experiment 1b, Chinese characters as stimuli. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean.

Figure 4. Proportion correct in visual working memory (VWM) tasks as a function of response modes, trial types, and stimulus types in Experiment 1a and 1b. Top panels, change detection; bottom panels, identity and location recall. Left-hand panels, Experiment 1a, English consonants as stimuli; right-hand panels, Experiment 1b, Chinese characters as stimuli. The error bars indicate standard errors of the mean.

Figure 5. Correlation between K_2 and Cowan's K in Experiment 1a and 1b. This analysis was based on both the non-rational strategy model and the all-or-none assumption. K_2 was averaged across array sizes and across identity and location recall tasks. Cowan's K was also averaged across array sizes from the change detection task. Left-hand panel, Experiment 1a with English consonants as stimuli; right-hand panel, Experiment 1b with Chinese characters as stimuli.

Figure 6. Diagram of a multinomial model of performance in the change detection task. With the all-or-none assumption for English consonants, Y=0. In the second branch (i.e., Critical Item→ In VWM → Low Resolution), X_1 is for hits and X_2 is for correct rejections. In the third branch (i.e., Critical Item→ Not in VWM), g is for hits and (1-g) is for correct rejections. Within the three branches, respectively, errors occur with probability 0, (1-X_1) or (1-X_2) for hits or correct rejections, and (1-g) or g for hits or correct rejections.

Figure 7. Diagram of a multinomial model of performance in the identity recall task. It is based on the non-rational model. With the all-or-none assumption for English consonants, Y=0. In the second branch (i.e., Critical Item→ In VWM → Low Resolution), X_3 is for accurate responses. In the third branch (i.e., Critical Item→ Not in VWM), $1/N$ is for accurate responses. Within the three branches, respectively, errors occur with probability 0, (1-X_3), and ($1-1/N$).

Figure 8. Proportion correct (solid points) along with predictions of the models of Figures 6 and 7 (open points). Parameter values in the model were selected by trial and error. Top panels, change detection; bottom panels, identity recall. Left-hand panels, Experiment 1a, English consonants as stimuli; right-hand panels, Experiment 1b, Chinese characters as stimuli. Ht = hit rate in the change detection task, Cr = correct rejection rate in the change detection task, and Ac = accuracy (proportion correct) in the identity recall task.

Figure 9. Individual residuals from the fitting of the partial information model, plotted against trial types, array sizes and stimulus types. Top panels, Experiment 1a with English consonants as stimuli; Bottom panels, Experiment 1b with Chinese characters as stimuli. Left-hand panels, individual residuals from predicted mean hit rates; Middle panels, individual residuals from predicted mean correct rejection rate; Left-hand panels, individual residuals from predicted mean identity recall accuracy.

Figure 10. Schematic illustration of procedure in Experiment 2.

Figure 11. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) in Experiment 2, which involved English consonants as stimuli and was under articulatory suppression. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean. Notice that VWM capacity estimates diverged at larger set sizes, unlike the left-hand panel of Figure 3 (English consonants, no

suppression) but more like the right-hand panel of Figure 3 (Chinese characters, no suppression).

Figure 12. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) as a function of the sample array display duration in Experiment 3, which involved Chinese characters as stimuli. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean. Notice that despite the use of an all-or-none assumption, the three capacity estimates approach equality for the longer sample display duration.

Figure 1

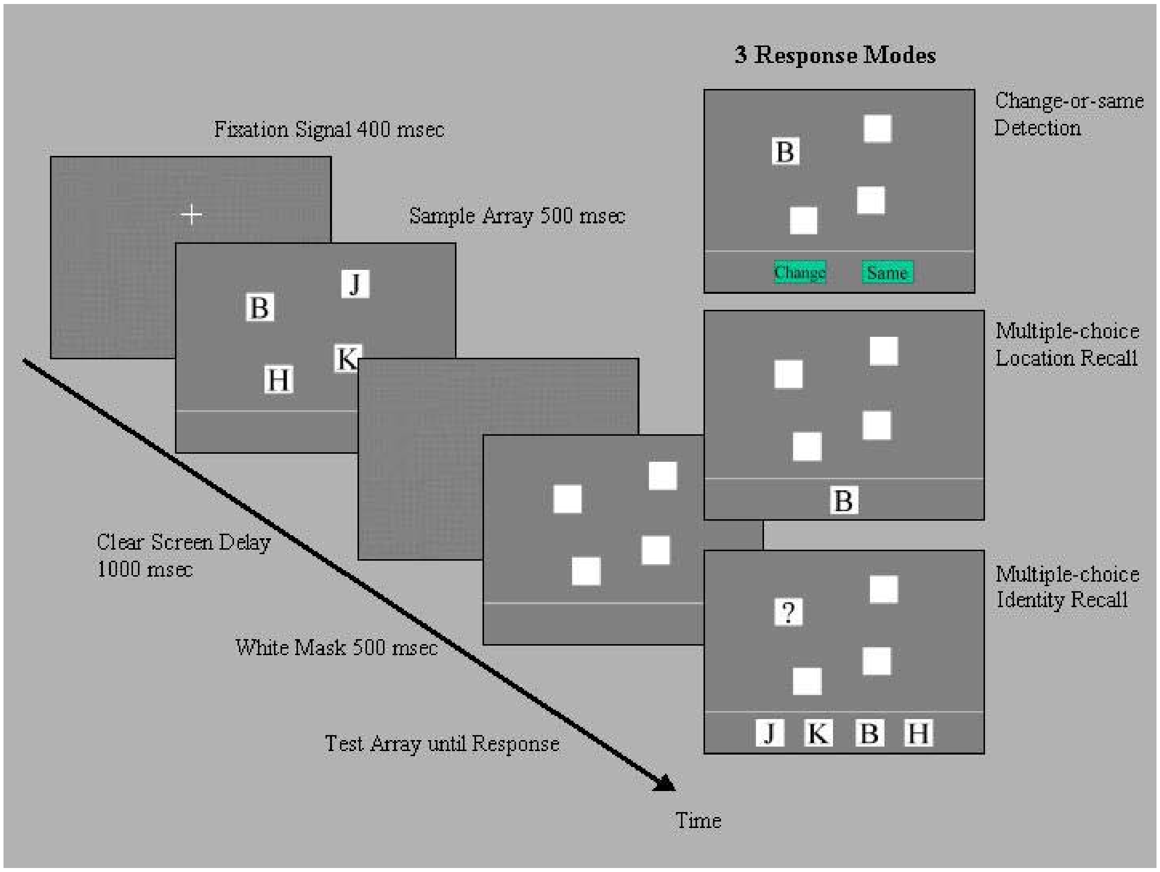


Figure 1. Schematic illustration of procedure in Experiment 1a. Experiment 1b shared the same procedure but with Chinese characters as stimuli.

Figure 2

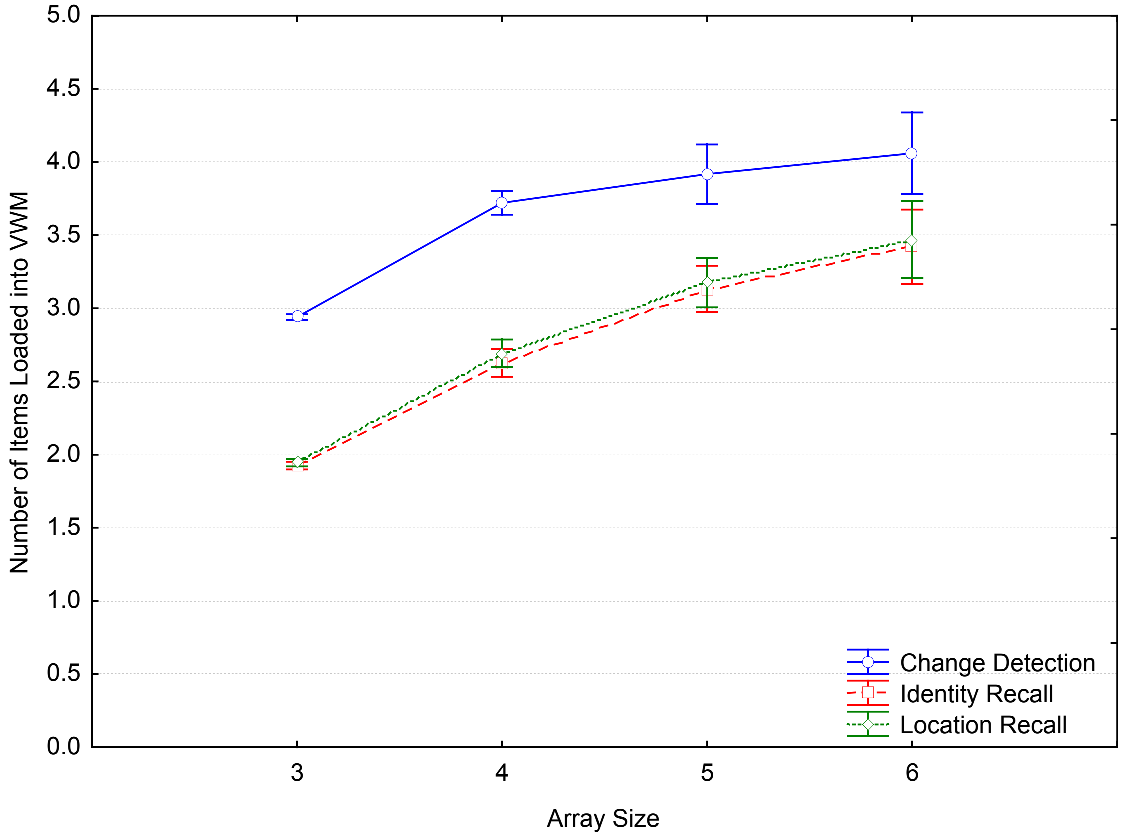


Figure 2. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_1 (for identity and location recall) as a function of response modes in Experiment 1a with English consonants as stimuli. This analysis was based on the all-or-none assumption and the "rational" strategy model. The error bars indicate standard errors of the mean.

Figure 3

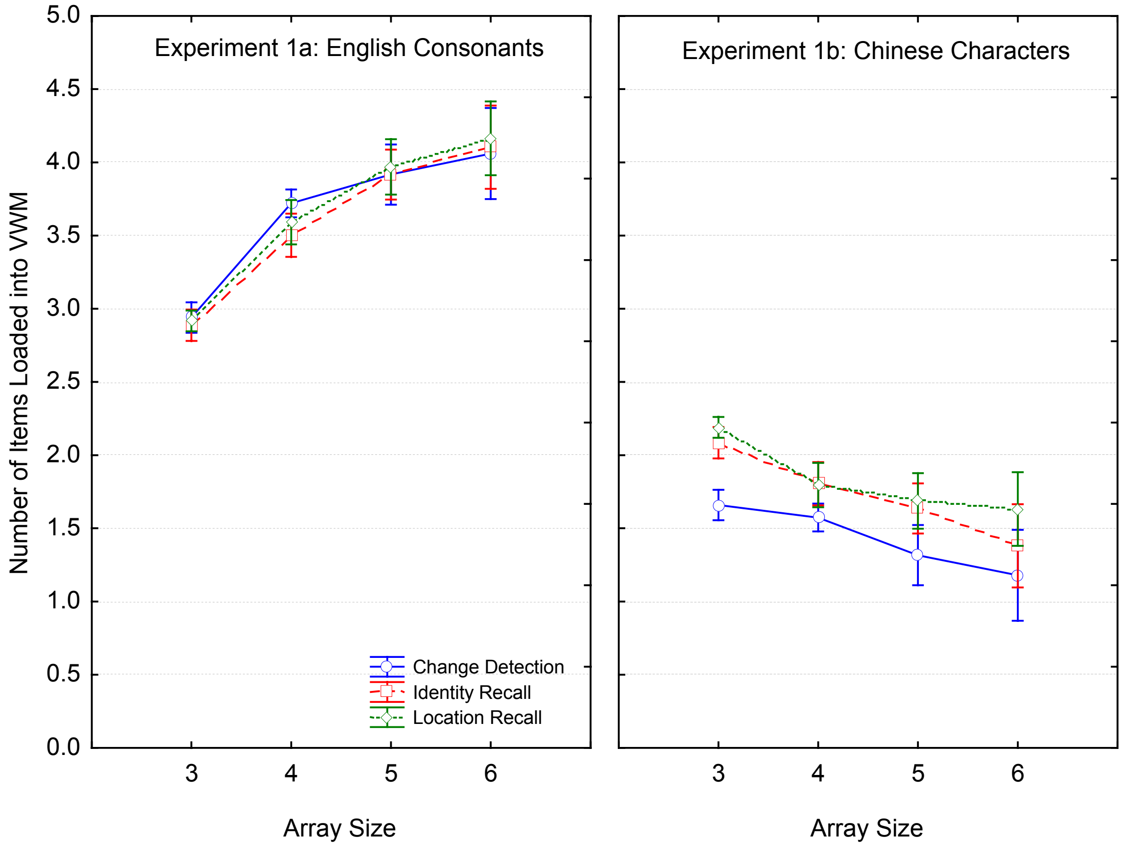
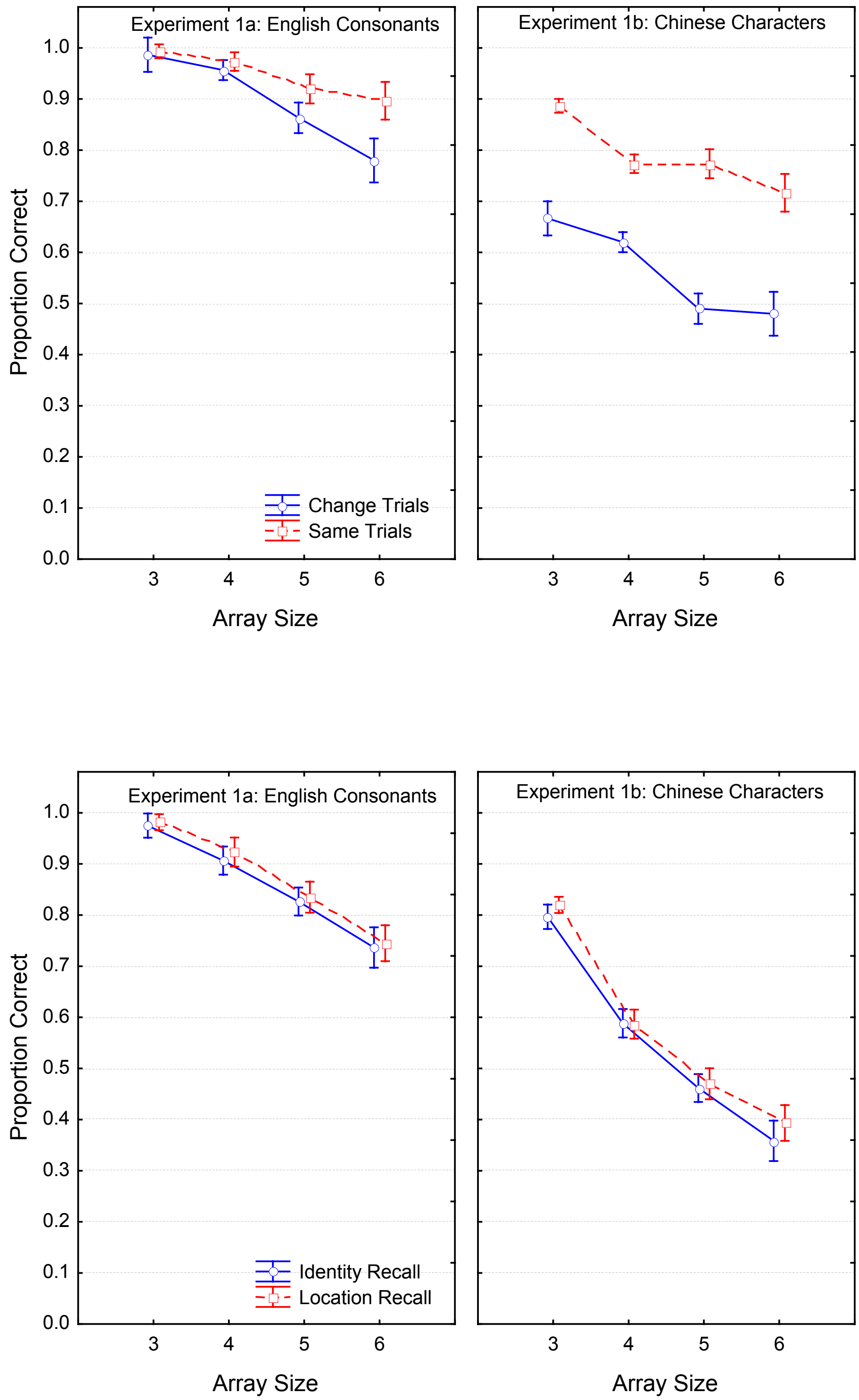


Figure 3. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) as a function of response modes. Left, Experiment 1a, English consonants as stimuli; right, Experiment 1b, Chinese characters as stimuli. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean.

Figure 4



Experiment 1a: English Consonants
Experiment 1b: Chinese Characters
Proportion Correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Change Trials
Same Trials
3
4
5
6
Array Size
Experiment 1a: English Consonants
Experiment 1b: Chinese Characters
Proportion Correct
Identity Recall
Location Recall
Array Size

Figure 4. Proportion correct in visual working memory (VWM) tasks as a function of response modes, trial types, and stimulus types in Experiment 1a and 1b. Top panels, change detection; bottom panels, identity and location recall. Left-hand panels, Experiment 1a, English consonants as stimuli; right panels, Experiment 1b, Chinese characters as stimuli. The error bars indicate standard errors of the mean.

Figure 5

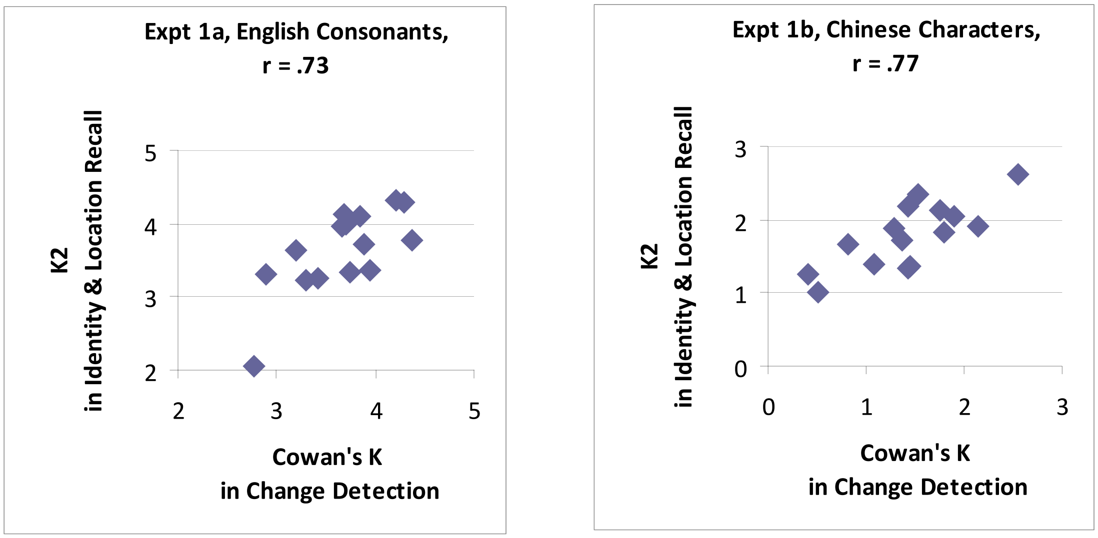


Figure 5. Correlation between K_2 and Cowan's K in Experiment 1a and 1b. This analysis was based on both the non-rational strategy model and the all-or-none assumption. K_2 was averaged across array sizes and across identity and location recall tasks. Cowan's K was also averaged across array sizes from the change detection task. Left-hand panel, Experiment 1a with English consonants as stimuli; right-hand panel, Experiment 1b with Chinese characters as stimuli.

Figure 6

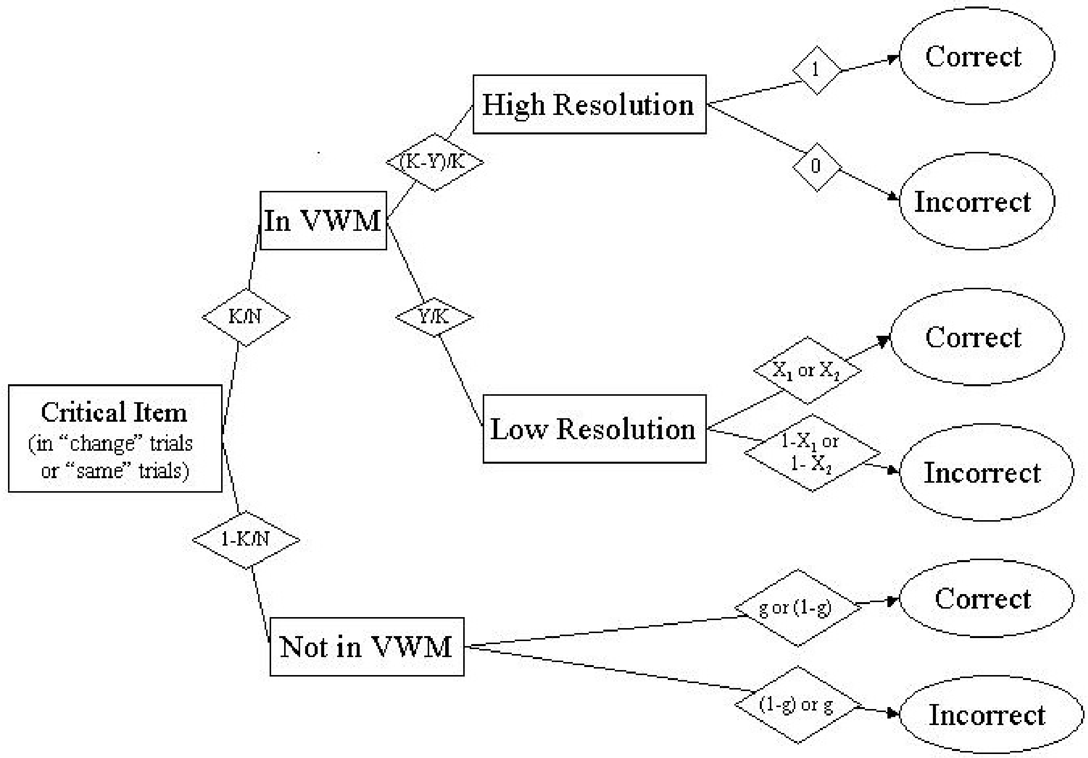


Figure 6. Diagram of a multinomial model of performance in the change detection task. With the all-or-none assumption for English consonants, Y=0. In the second branch (i.e., Critical Item→ In VWM → Low Resolution), X_1 is for hits and X_2 is for correct rejections. In the third branch (i.e., Critical Item→ Not in VWM), g is for hits and (1-g) is for correct rejections. Within the three branches, respectively, errors occur with probability 0, (1-X_1) or (1-X_2) for hits or correct rejections, and (1-g) or g for hits or correct rejections.

Figure 7

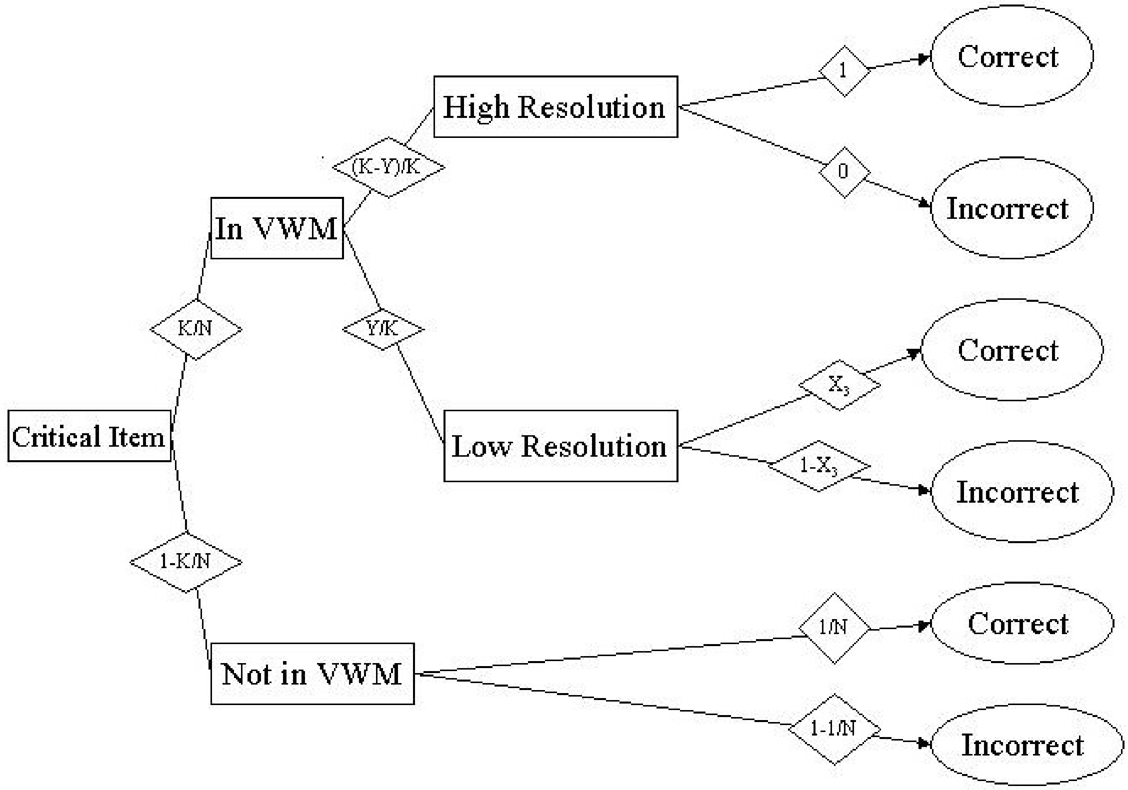


Figure 7. Diagram of a multinomial model of performance in the identity recall task. It is based on the non-rational model. With the all-or-none assumption for English consonants, Y=0. In the second branch (i.e., Critical Item→ In VWM → Low Resolution), X_3 is for accurate responses. In the third branch (i.e., Critical Item→ Not in VWM), *1/N* is for accurate responses. Within the three branches, respectively, errors occur with probability 0, (1-X_3), and (*1-1/N*).

Figure 8

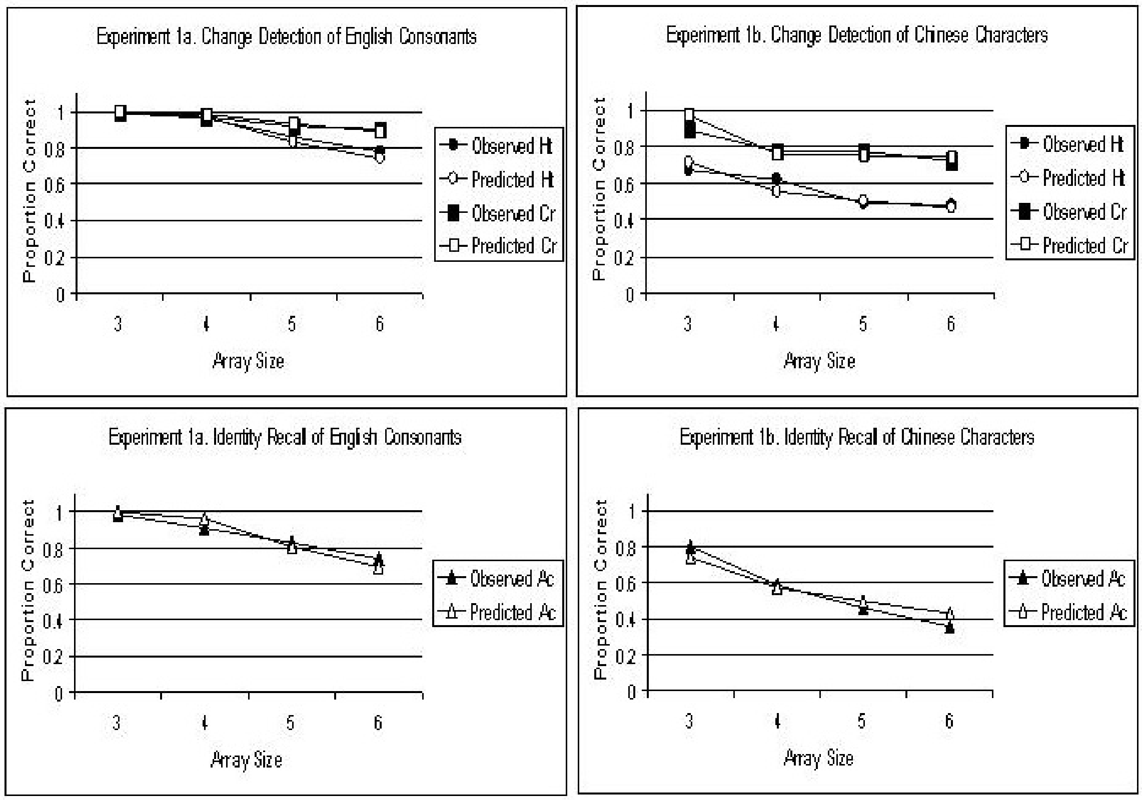


Figure 8. Proportion correct (solid points) along with predictions of the models of Figures 6 and 7 (open points). Parameter values in the model were selected by trial and error. Top panels, change detection; bottom panels, identity recall. Left-hand panels, Experiment 1a, English consonants as stimuli; right-hand panels, Experiment 1b, Chinese characters as stimuli. Ht = hit rate in the change detection task, Cr = correct rejection rate in the change detection task, and Ac = accuracy (proportion correct) in the identity recall task.

Figure 9

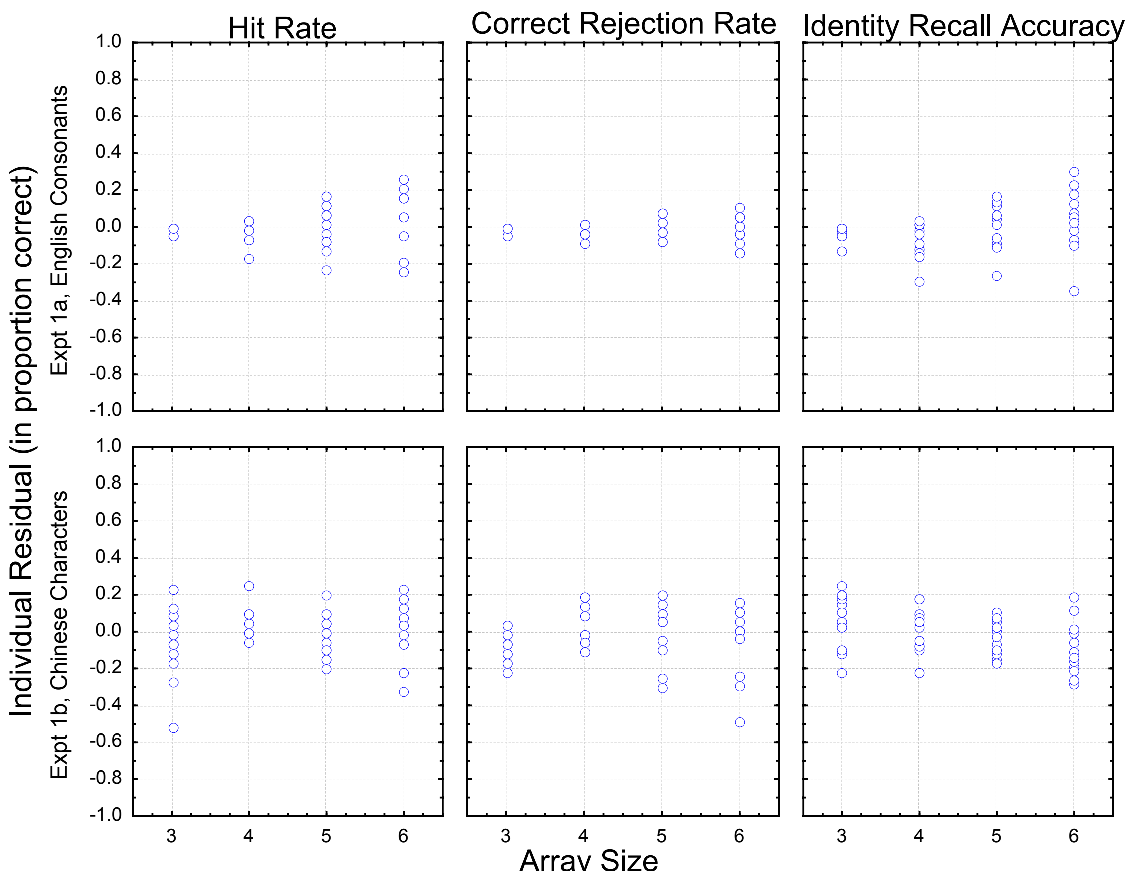


Figure 9. Individual residuals from the fitting of the partial information model, plotted against trial types, array sizes and stimulus types. Top panels, Experiment 1a with English consonants as stimuli; Bottom panels, Experiment 1b with Chinese characters as stimuli. Left-hand panels, individual residuals from predicted mean hit rates; Middle panels, individual residuals from predicted mean correct rejection rate; Left-hand panels, individual residuals from predicted mean identity recall accuracy.

Figure 10

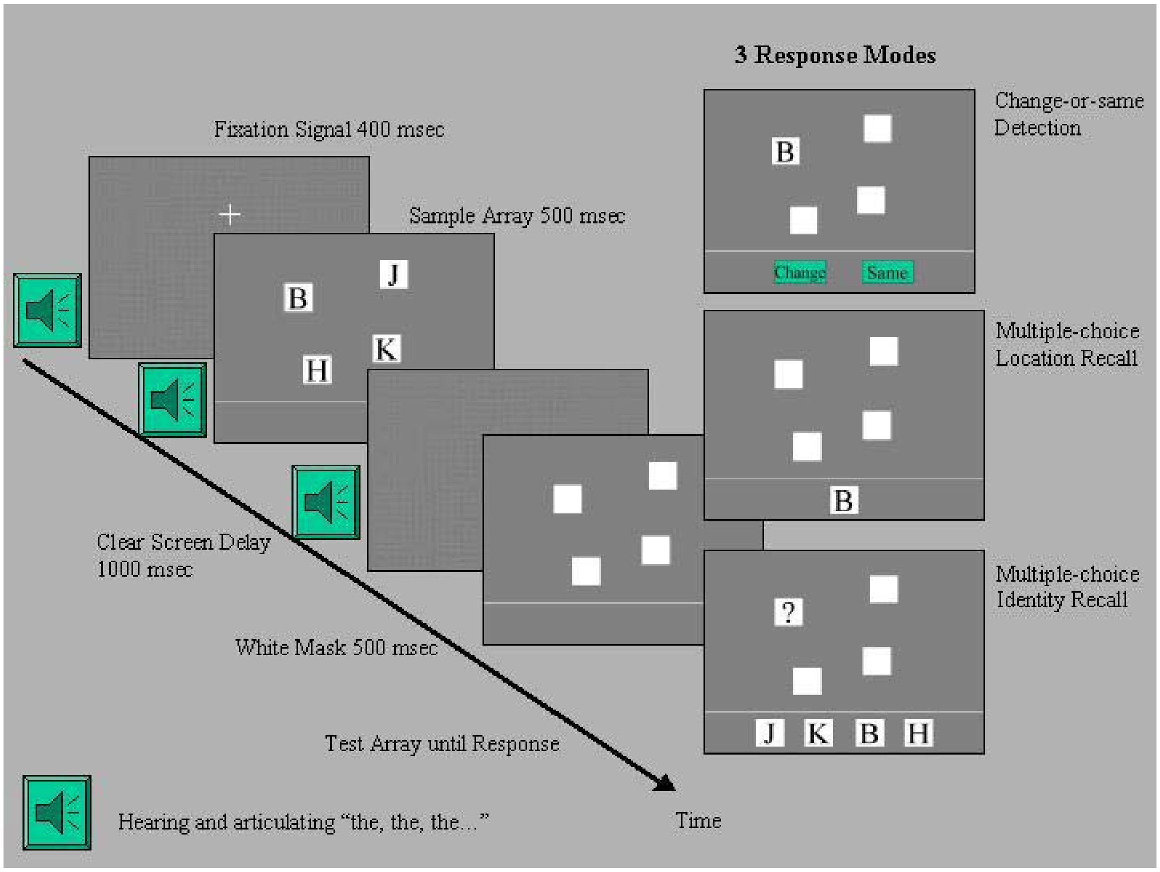


Figure 10. Schematic illustration of procedure in Experiment 2.

Figure 11

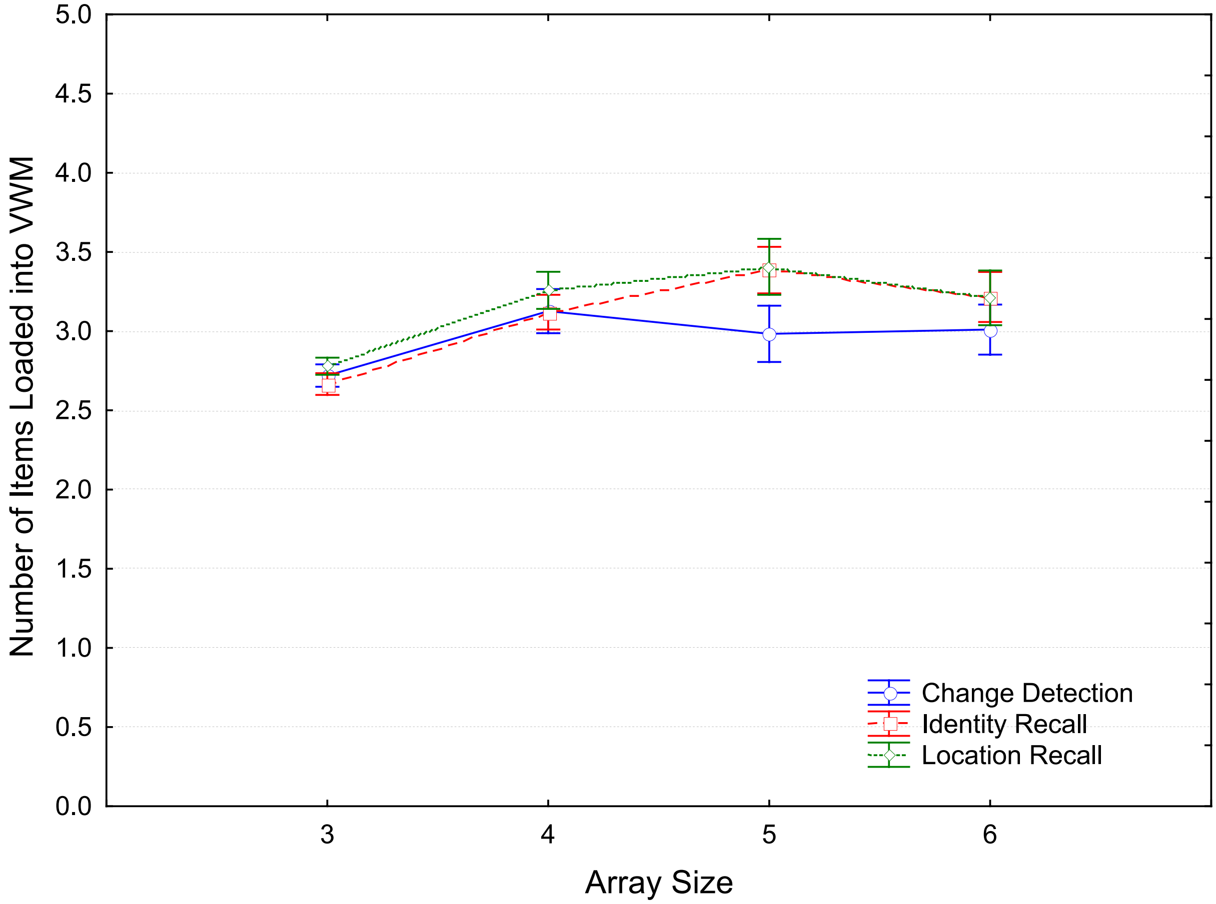


Figure 11. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) in Experiment 2, which involved English consonants as stimuli and was under articulatory suppression. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean. Notice that VWM capacity estimates diverged at larger set sizes, unlike the left-hand panel of Figure 3 (English consonants, no suppression) but more like the right-hand panel of Figure 3 (Chinese characters, no suppression).

Figure 12

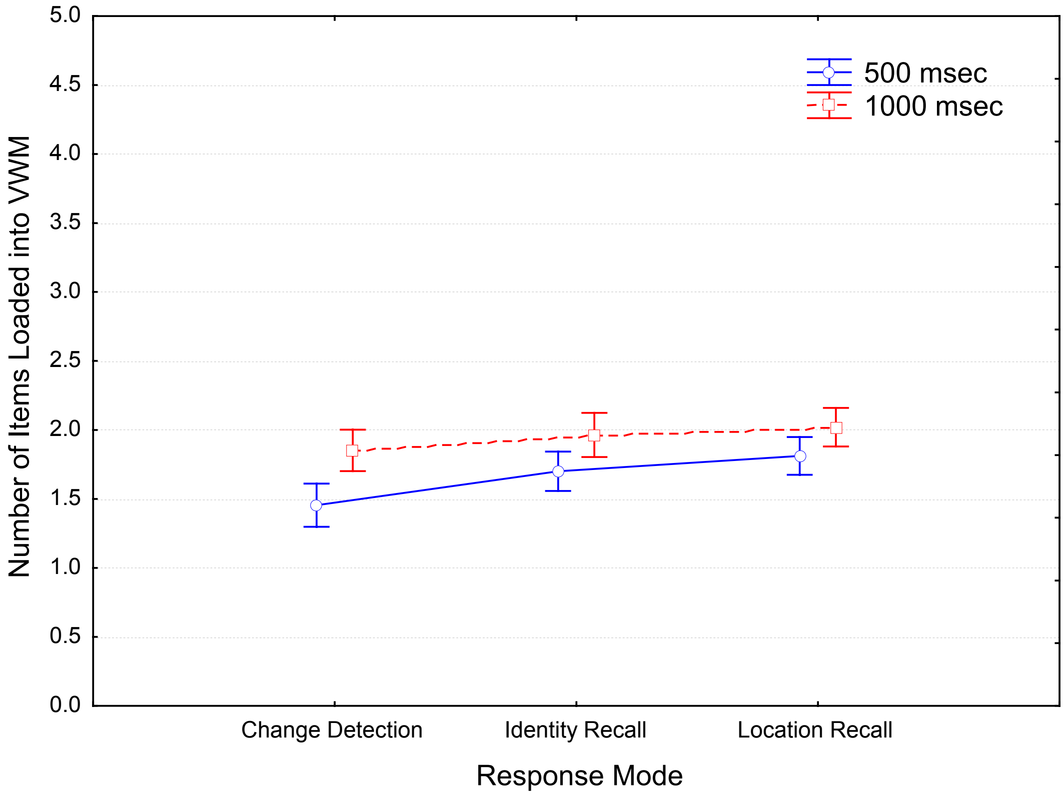


Figure 12. Visual working memory (VWM) capacity estimates, Cowan's K (for change detection) and K_2 (for identity and location recall) as a function of the sample array display duration in Experiment 3, which involved Chinese characters as stimuli. This analysis was based on the all-or-none assumption and the "non-rational" strategy model. The error bars indicate standard errors of the mean. Notice that despite the use of an all-or-none assumption, the three capacity estimates approach equality for the longer sample display duration.

VITA

Zhijian (David) Chen was born in Guangzhou, P.R. China on January 8, 1972, the son of Hao Chen and Xuemei Guo. After completing high school at the Affiliated High School of South China Normal University in 1991, he attended Zhongshan University from 1991 to 1995. He graduated with a Bachelor of Arts degree in Economics in 1995. He then entered Renmin University of China in the fall of 1995 and received his Master of Arts degree in Political Economics in 1998.

David started to pursue a new academic career in cognitive psychology at University of Missouri-Columbia, U.S.A. in 2001. He attained his Master of Arts degree in Psychology in 2004 and Doctor of Philosophy degree in Psychology in 2009 from the University of Missouri-Columbia under the supervision of Dr. Nelson Cowan.