

Removal of Information from Working Memory: A Specific Updating Process

Ullrich K. H. Ecker

School of Psychology, University of Western Australia

Stephan Lewandowsky

School of Experimental Psychology, University of Bristol, and

School of Psychology, University of Western Australia

Klaus Oberauer

Department of Psychology, University of Zurich, and

School of Psychology, University of Western Australia

Author Note

Preparation of this paper was facilitated by Discovery Grants and Research Fellowships from the Australian Research Council to Ullrich Ecker and Stephan Lewandowsky. We thank Charles Hanich and Devon Spaapen for research assistance, and Yoav Kessler and Nelson Cowan for their comments on earlier drafts. Address correspondence to the first author at the School of Psychology (M304), University of Western Australia, Crawley 6009, Australia; e-mail: ullrich.ecker@uwa.edu.au. The lab's web site is located at www.cogsciwa.com.

Word count: 7,570 excluding abstract, references, and figure captions

Abstract

Previous research has claimed that working memory (WM) updating is one of three primary central executive processes, and the only one to reliably predict fluid intelligence. However, standard WM updating tasks confound updating requirements with generic WM functions. This article introduces a method for isolating a process unique to WM updating, namely the removal of no-longer relevant information. In a modified version of an established updating paradigm, to-be-updated items were cued before the new memoranda were presented. Overall, longer cue-target intervals—that is, longer time available for removal of outdated information—led to faster updating, suggesting that people can actively remove information from WM. Experiments 1 and 2 demonstrated that well-established effects of item repetition and similarity on updating RTs were diminished with longer cue-target interval, arguably because representational overlap between outdated and new information becomes less influential when outdated information can be removed prior to new encoding. Experiment 3 looked at individual differences, using the reduction of updating RTs to measure removal speed. Removal speed was measured reliably but was uncorrelated to WM capacity. We conclude that (1) removal of outdated information can be experimentally isolated and measured reliably, (2) removal speed is a unique, active WM updating ability, and (3) the view of WM updating as a core executive process that uniquely predicts fluid abilities is overstated.

Keywords: working memory; updating; removal; individual differences; executive functions;

SOB

Removal of Information from Working Memory: A Specific Updating Process

Imagine you ask a colleague for his phone extension and he replies: “It’s 3266. No, hang on, in my new office it’s actually 3257”. Ideally, one should easily discard the last two digits of the outdated information given (i.e., “66”) and replace them in working memory with the correct digits (i.e., “57”). However, this updating of working memory content is no trivial task, and outdated information often continues to affect memory (De Beni & Palladino, 2004; Oberauer, 2001).

Working memory updating has been identified as one of three primary central executive processes (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Updating has been claimed to be the only executive process to predict fluid intelligence (Chen & Li, 2007; Friedman et al., 2006). However, most updating tasks used in previous research (e.g., Miyake et al., 2000) not only require memory updating but arguably also measure general working memory (WM) abilities. This has led some researchers to conclude that updating tasks constitute reliable assays of general WM capacity (Schmiedek, Hildebrandt, Lovdén, Wilhelm, & Lindenberger, 2009; see also Chuderski, Taraday, Nęcka, & Smoleń, 2012; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Martínez et al., 2011).

This creates an unsatisfactory situation. If WM updating tasks measure just the same as other WM tasks such as complex span tasks, then there is no empirical basis for identifying ‘updating’ as a separate executive-function factor. Yet, both conceptually and theoretically, updating can be distinguished from maintenance and processing in WM. If updating is to be established as a non-redundant construct, it must be isolated and measured separately from other WM processes.

In a recent individual-differences study, we identified a processing component that was independent of general WM capacity and unique to situations that demanded memory updating (Ecker, Lewandowsky, Oberauer, & Chee, 2010). In that study we analyzed the processing components involved in widely used WM updating tasks, and we identified three separable components: retrieval, transformation, and substitution. The only component process that was unique to WM updating tasks was the substitution of information in memory. To illustrate those components, consider the scenario of a restaurant manager advising a chef early in the evening that they were expecting 20 patrons. If the manager later advised the chef that twice as many guests were expected as before, the chef will need to retrieve the initial expectation (i.e., 20), transform it (i.e., $2 \times 20 = 40$), and substitute the outdated information with the updated information (i.e., 40). Ecker et al. designed an updating task with eight conditions, fully crossing all possible combinations of retrieval, transformation, and substitution. Applying structural equation modeling to their data, they found that retrieval and transformation operations co-varied with general WM capacity, but that the substitution component did not. This is illustrated by the structural equation model for their updating accuracy data, shown in Figure 1. This finding was interpreted as showing that substitution is the only process that uniquely represents WM updating, without being “contaminated” by any association with general working memory abilities.

One implication of this analysis is that previous studies measuring WM updating did not separate variance unique to updating from the variance of generic WM processes. As a consequence, the conclusions concerning the predictive relation between WM updating and fluid intelligence (Chen & Li, 2007; Friedman et al., 2006) were arguably not based on a proper measure of WM updating, but may instead reflect the well-known association between higher

cognitive functions and general WM capacity (Engle, Tuholski, Laughlin, & Conway, 1999; Oberauer, Schulze, Wilhelm, & Süß, 2005).

In this article, we further decompose the components of WM updating. In Ecker et al. (2010), we suggested that information substitution can be further subdivided into the removal of outdated information and the encoding of new information. For example, the chef would need to remove the number 20 from memory and encode the updated number 40 into the vacant memory slot. Of these two processes, encoding is common to many cognitive tasks. In contrast, we argue that it is the removal process that lies at the heart of, and is specific to, memory updating. Accordingly, we focus on the removal of information from WM. Specifically, we define removal as the unlearning or unbinding of an item from its context. Here we show that (1) removal of outdated information can be separated experimentally from encoding of new information, (2) that removal can be characterized as an active WM updating process, (3) that removal speed can be measured reliably, and (4) that removal speed is independent of WM capacity.

Our conceptualization of removal as an active WM process is inspired by a computational model of working memory, the SOB (“serial-order in a box”) model (Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). SOB is a two-layer neural network in which items (represented in one layer) are associated to position or context markers (represented in the other layer) through Hebbian learning, which rapidly modifies the matrix of connection weights between the two layers. Memory for items-in-position is hence stored in this weight matrix. Forgetting in SOB is entirely based on interference; there is no time-based trace decay. This assumption is corroborated by a growing body of evidence (Berman, Jonides, & Lewis, 2009; Jalbert, Neath, Bireta, & Surprenant, 2011; Lewandowsky, Oberauer, & Brown, 2009; Oberauer & Lewandowsky, 2008, 2013).

To avoid overloading of the system in the absence of decay, an interference model requires a mechanism to remove outdated information; such a mechanism is implemented in the most recent version of SOB (SOB-CS, a model for the complex span task; see Oberauer et al., 2012). Removal of a specific item involves retrieving that item by cueing with its position marker, and “unlearning” the association between that item and its position. Unlearning is computationally implemented as Hebbian anti-learning. Whereas encoding of information into WM takes less than 500 ms per item (Jolicoeur & Dell’Acqua, 1998; Vogel, Woodman, & Luck, 2006), removal appears to be a slower process, taking at least about 500-600 ms per item. This estimate is taken from work using a directed-forgetting approach, where participants study two sets of items, one of which is then declared irrelevant before the relevant set is tested. The size of the irrelevant set continues to affect responses to the relevant set for about 1-2 seconds when the largest irrelevant set comprises three items (Oberauer, 2001; see also LaRocque, Lewis-Peacock, Drysdale, Oberauer, & Postle, 2013). Hence, the time it takes to remove a single item can be estimated as roughly a third of that time, if removal of multiple items occurs sequentially.

Our removal measure is based on the work of Kessler and Meiran (2008), who used a variant of the classic updating paradigm developed by Yntema and Mueser (1962). In this paradigm, individual items (e.g., letters or digits) are presented in a set of individual frames. Items are then repeatedly updated by presenting new items in some frames. On each updating step, at least one item is updated, and participants have to press a key at the end of each step to indicate that they finished updating; this provides an updating RT measure. At the end of the sequence, participants recall the currently memorized set of items.

In SOB, removal of old information and encoding of new information are described as two separate processing steps. It follows that updating should be facilitated if a cue about what

information needs to be removed is given ahead of the to-be-encoded new information. In contrast, an incidental feature of standard WM updating tasks to date—including the one used by Kessler and Meiran (2008)—has been that removal cannot commence until the new information is presented for encoding. This is due to the way standard WM updating tasks are designed, namely that the time-point at which people are told what to update coincides with the time-point at which they are given the new information to encode. For example, when the current set of letters in a 3-frame updating task is ‘K-M-R’ and a new letter ‘D’ appears in the third frame, removal of the ‘R’ can only begin when the to-be-updated frame is identified, that is, when the ‘D’ is displayed. Hence updating response times in such a task will include both time for removal and time for encoding. This confound can be avoided by signaling which frame will be updated before the new information is presented. In our new updating task, we thus present cues indicating which items are to be updated before presenting the new to-be-encoded stimuli. People can use the cue only to selectively remove old items from the memory set, not to encode new information. By varying the cue-target interval (CTI) we vary the available time for removal. If people use the CTI to remove the outdated information in the cued frame, then longer CTIs should lead to faster updating RTs.

In Experiments 1 and 2 we show that this is the case, and argue that the reduction of updating RTs by longer CTIs in fact reflects the efficiency of removal during the CTI. Specifically, Experiments 1 and 2 supply an experimental validation of our proposed measure of removal efficiency by (1) showing that there is a big updating RT gain from removing in advance (i.e., a long CTI), and by (2) ruling out alternative explanations for this finding by testing a prediction that is unique to the removal notion, namely that benefits associated with item repetition (Exp. 1) and item similarity (Exp. 2) diminish with a long CTI. This will be followed

by Experiment 3, which tests the main hypothesis of our paper, namely that removal efficiency—the one process that is most specific to memory updating—is uncorrelated with WM capacity.

Our first investigation of the effect of the CTI on updating RTs involved testing a specific prediction derived from Ecker et al. (2010). One of the Ecker et al. conditions involved the straightforward substitution of an item; for example, while remembering ‘K-M-R’ in three frames, a ‘D’ might have been presented in the third frame, prompting the participant to update to ‘K-M-D’ (this is the $r_{no}-t_{no}-S$ condition in Figure 1). Another condition was a control condition where an item was presented repeatedly in the same frame on successive updating steps. For example, while a participant was remembering the letters ‘K-M-R’, an ‘M’ might have been presented in the middle frame, prompting a participant to continue to remember ‘K-M-R’ without any need to update the memory representation (this is the $r_{no}-t_{no}-s_{no}$ condition in Figure 1). We found that repeating an item during the updating task (i.e., maintaining an item unchanged) carried a benefit of roughly 300-400 ms. This benefit should diminish when people are given the opportunity to remove outdated information before encoding the (in this case, identical) updated item, that is, with a long CTI.

Experiment 1

Experiment 1 used a letter updating task in which each trial consisted of an encoding stage, an updating stage with multiple updating steps, and a final recall stage. Participants encoded 3 letters, presented simultaneously in individual frames. This was succeeded by an unpredictable number of updating steps; each updating step involved only a single, randomly selected letter. In most cases, the outdated letter was replaced with a new letter, but sometimes the letter was identical to the previous letter presented in the same frame (and was thus maintained unchanged in WM). Before each updating step, participants were warned which item was about to

be updated; to this end, each update was cued before the new letter was presented (the respective frame turned red and bold). Presentation time of this cue—the cue-target interval—was either short or long, the rationale being that the longer cue should provide time for removal, whereas the shorter cue should not. The experiment had a 2 (repetition no/yes) \times 2 (CTI: short/long) within-subjects design.

Methods

Participants. We tested 15 participants, mainly undergraduate students from the University of Western Australia (8 females, 7 males; mean age was 23.8 years, age range 17-29 years).

Apparatus. The experiment was run with the aid of MatLab and the Psychophysics toolbox (Brainard, 1997). Participants were tested individually in booths, sitting about 70 cm from a 17-in. thin-film transistor monitor.

Stimuli, Design, and Procedure. The letter updating task involved three letters presented in a single row of three black, rectangular frames. A representative (albeit short) trial sequence is shown in Figure 2. Following a fixation cross presented for one second, each trial began with the simultaneous 2-second presentation of 3 black consonants (ranging from B to Z) in the frames. The minimum alphabetic distance between the to-be-encoded letters was 2, to avoid sequences such as X, Y, Z.

A series of updating steps followed. On each step, one of the letters was cued for removal, with either a short or a long cue-target interval, and a new letter was presented in the corresponding frame. Figure 3 gives a schematic depiction of an updating step for both the long and the short CTI condition.

Specifically, each updating step consisted of three phases: (1) a response-cue interval during which only the empty frames were visible, (2) a cue-target interval (CTI) during which one of the frames turned bold and red, signaling to participants that the letter remembered in that position was about to be updated, and (3) the target-display phase, during which the new letter was presented until participants pressed the space bar to indicate that they had completed the updating step (i.e., that they had encoded the new content and updated their WM), with a maximum response time of 5 seconds. This updating RT was the main dependent variable of interest. The CTI was either short (i.e., 200 ms before presentation of the new letter) or long (i.e., 1500 ms). The longer cue should be sufficient time for removal, whereas the shorter cue should be just enough time to focus attention on the to-be-updated frames without permitting removal. The CTI was chosen randomly at each step. The empty-frames response-cue interval was chosen to be complementary to the CTI, such that they added to a constant 2-second retention interval in both conditions. Thus, the response-cue interval was 500 ms in the long CTI condition and 1800 ms in the short CTI condition.

The number of updating steps per trial varied from 1 to 21; the sequence finished with a constant probability of 10% after each updating step. This resulted in an unpredictable number of updating steps, and because each step was equally likely to be the last, participants had an equal incentive to carry out each updating step independent of the duration of the trial. Items repeated with a probability of 15%; this rate was kept low to ensure that participants still perceived the removal of cued information as an attractive strategy. Repetitions could be immediate (e.g., encoding a 'B' in the middle frame on two successive updating steps), or separated by an unspecified number of updates in other frames (e.g., encoding a 'B' in the middle frame, then updating frame 1 and/or frame 3, then encoding a 'B' in the middle frame again).

After the updating phase of each trial, participants recalled the current contents of all frames. Recall was prompted by blue question marks appearing one-by-one in each frame in random order. Probes remained on the screen until a response was given, or until the maximum response time of 5 seconds per frame was reached. After recall of all three frames, feedback (“x out of 3 correct”) was given. The blank-screen inter-trial interval was 2.5 seconds.

There were 32 trials in total (plus 2 practice trials), with an average of 9 updating steps per trial, yielding approximately 288 updating steps (about 22 per CTI condition when items repeated, and 122 per condition when items did not repeat). Each trial took approximately 40 seconds, and the experiment took about 25 minutes.

Results

Recall accuracy was very high ($M = 0.97$, $SE = .009$). In all experiments, updating RTs below 300 ms, and more than 3 standard deviations from the individual mean were discarded. Updating RT data are shown in Figure 4.

A 2×2 repeated measures ANOVA on updating RTs with the factors repetition (no/yes) and removal time (short/long) yielded a main effect of repetition, $F(1, 14) = 15.23$, $MSE = 0.03$, $p < .01$, $\eta_p^2 = 0.52$, a main effect of removal time, $F(1, 14) = 9.78$, $MSE = 0.01$, $p < .01$, $\eta_p^2 = 0.41$, and crucially, a significant interaction, $F(1, 14) = 8.22$, $MSE = 0.01$, $p = .01$, $\eta_p^2 = 0.37$. The interaction indicates that the effect of repetition on updating RTs was larger with short removal (265 ms) than long removal time (86 ms; contrast analysis showed this was still a significant difference, $F(1,14) = 5.70$, $MSE = 0.01$, $p = .03$).

Discussion

Experiment 1 demonstrated that the benefit of item repetition during memory updating (Ecker et al., 2010) is strongly reduced if participants are given time to remove outdated

information prior to the encoding of the updated information. In a sense, if there is no time for removal, a condition with repeating items does not require true memory updating—the established item representation in WM can be maintained, no information needs to be removed and substituted—hence the time advantage of repetition. In contrast, to the degree that an item is removed from WM, the time taken to encode a new item into that position will no longer depend on the identity of the removed item. In other words, to the degree that an item is removed from WM, the benefit of repetition vanishes because the old item is no longer represented in WM. The obtained result pattern thus supports our notion that active removal of information is integral to WM updating.

The updating RT difference between the short and long CTI conditions with no repetition—nearly 200 ms—can be interpreted as the minimum time required to initiate removal, and achieve some level of significant but partial removal of information from working memory. We do not argue that the removal process is completed within 200 ms. Previous research has estimated near-complete removal to take at least 500-600 ms per item (e.g., LaRocque et al., 2013; Oberauer, 2001), and complete removal to take up to 2 seconds (Oberauer et al., 2012). In fact, the rudimentary repetition effect with a long CTI indicates that removal in the present experiment was not fully completed even after 1500 ms (or alternatively, that removal occasionally failed, or some mixture of these two possibilities); otherwise, the long-CTI repetition trials would be as slow as the long-CTI no-repetition trials. Moreover, the 200 ms figure could be an underestimation, to the degree that (in the short CTI condition), removal and encoding can partially run in parallel—it seems reasonable to assume that early perceptual encoding processes at least can progress while removal of outdated information is ongoing.

At first glance our rationale would seem to imply that if items repeat, then RTs should be slower in the long compared to the short CTI condition: The more an about-to-be-repeated item is removed, the more time should be required to re-encode it. However, this line of argument holds only if the processes in repetition trials with long and with short CTI are directly comparable. We argue that they are not: In repetition trials with long CTI, participants largely remove the old representation during the CTI, and therefore, the RTs reflect primarily the time for encoding the new item. In repetition trials with short CTI, participants don't remove the old item, and their RTs reflect the time for detecting the identity between the old and the new item, for stopping the default updating process, and perhaps for carrying out a refreshing operation on the old (and still current) representation in WM instead. Therefore, the fact that the updating RTs in these two conditions are quite similar does not imply that they reflect the same underlying processes (for further investigations relevant to this issue, see Ecker, Lewandowsky, & Oberauer, 2013; Kessler & Oberauer, 2013).

There are, however, two potential alternative accounts of the data, which we will discuss in turn. First, the long CTI could lead to a greater general 'preparedness' or alertness, which would also lead to a speed-up. A general alertness account would, however, predict that RTs in both the repetition and non-repetition conditions should be reduced to about the same degree with a long CTI, which was not the case. Thus, to explain the present data pattern, one would have to assume a more selective type of preparedness, namely a preparedness specifically to update WM content, which applies only to the non-repetition condition. We will address this issue in the context of Experiment 2.

Second, because of the need to equate the retention interval across conditions, our cue-target interval is fully confounded with the blank-frames interval preceding each update: When

the CTI is long (1500 ms) then the blank-frames interval preceding it is short (500 ms) and vice versa (200 ms and 1800 ms, respectively). One could hence argue that the response time differences we find could be influenced by some sort of short-term consolidation process (Jolicoeur & Dell'Acqua, 1998). In particular, items updated after a long blank-frames response-cue interval and a short CTI may be more difficult to update because they have had more time to consolidate.

This possibility applies only in cases where the updated frame on step n was also updated on step $n - 1$, because only in those cases could the to-be-removed old item be consolidated in the immediately preceding response-cue interval. A straightforward way to investigate this is to compare the data pattern in cases where the updated frame (on updating step n) was also updated on step $n - 1$ with cases where the updated frame was not updated on the previous step. In the former case, an item's memory representation may still be consolidating when the updating cue appears, but in the latter case, consolidation should be complete by that time (or at least, the additional 1300 ms should not matter much anymore).

To this end, we classified all updates not involving an item repetition depending on whether or not the same frame was also updated on the previous step. We then ran a 2×2 repeated-measures ANOVA with the factors CTI (short vs. long) and frame-switch (yes/no). The analysis returned a main effect of CTI, $F(1, 14) = 10.18$, $MSE = 0.04$, $p < .01$, $\eta_p^2 = 0.42$, and a main effect of frame-switch, $F(1, 14) = 14.52$, $MSE = 0.01$, $p < .01$, $\eta_p^2 = 0.51$, but no interaction, $F < 1$.

This analysis demonstrates two things: (1) Updating the same frame twice in a row is quicker than updating a frame that was not updated on the immediately preceding step. This is most likely because updating the same frame twice in a row requires less attentional switching

between frames; the RT costs of attentional switching are well-documented (Garavan, 1998; Oberauer, 2002). More importantly, (2) the effect of CTI is as large in frame-switch trials as in frame-repetition trials. In frame-switch trials, the time for consolidation of the to-be-removed item (encoded several steps back) is unrelated to the time for removing it (i.e., the current CTI). Therefore, the effect of CTI cannot be attributed to the time for consolidating the to-be-removed item. We hence conclude that our main result is not affected by the confound of cue-target interval and “consolidation time.”

Experiment 2

Experiment 2 had a similar rationale to Experiment 1. Previous research (Lendinez, Pelegrina, & Lechuga, 2011) has shown that updating numbers is quicker when the new to-be-remembered number is similar to the outdated number. For example, updating from 21 to 22 takes less time than updating from 21 to 27. Experiment 2 tested whether this benefit would diminish if participants were given sufficient time to remove the outdated number before encoding the new number. A diminished similarity effect is predicted from the removal hypothesis, but it is not predicted by the notion of an updating-specific preparedness caused by a long CTI: In both the similar and the dissimilar condition participants need to update the content of WM, so any preparedness to update should affect both conditions equally.

Method

The task in Experiment 2 was very similar to Experiment 1, but it used two-digit numbers. During updating, about half the updates used similar and dissimilar numbers, respectively. Lendinez et al. (2011) manipulated the similarity of updated numbers using a relatively small distance manipulation, comparing distances of 1 or 2 with distances of 5 or 6. Because digit repetition is much more likely to occur with small distances of 1 or 2 than distances of 5 or 6, this

is a difficult manipulation to implement without confounding distance with digit repetition. Moreover, this manipulation does not utilize the full range of distance/similarity. For these reasons, we decided to deviate from the design of Lendinez et al.

To this end, we manipulated both the proximity of numbers (proximal/distant) and the repetition of one of the digits (yes/no; e.g., the numbers 18 and 19 share the digit '1' in the first position, the numbers 35 and 45 share the digit '5' in the second position; we only considered such in-position repetitions). This resulted in four updating conditions, three with some degree of similarity and one using dissimilar numbers: proximal/repeating (e.g., updating from 18 to 19), proximal/non-repeating (e.g., updating from 20 to 18), distant/repeating (e.g., updating from 59 to 19), and the dissimilar distant/non-repeating condition (e.g., updating from 18 to 55). Numeric distance between old (outdated) and new (updated) numbers ranged from -3 to +3 (excluding zero) in proximal updates; distant/repeating updates were constrained to distances that were multiples of 10; and distant/non-repeating updates used prime numbers from 13 to 83 as distances.¹ Because the a-priori plan was to contrast the three similar conditions with the dissimilar condition, the four conditions had different probabilities of occurrence, which were $p = .15$ for each of the three similar conditions, and $p = .50$ for the dissimilar condition. Filler updates with intermediate proximity (distances $\pm 4 - 8$) were used with $p = .05$. Updates replacing the previous number with a larger or smaller number (e.g., going from 18 to 19 or from 19 to 18, respectively) were randomly intermixed (this was not considered an experimental factor). CTI (short/long) was again an additional factor.

Participants. We tested 27 participants, all undergraduate students from the University of Western Australia.

Apparatus & Procedure. The apparatus and procedure was identical to Experiment 1, with the exception that the experiment had 36 trials.

Results

The task in Experiment 2 was somewhat more difficult than the task in Experiment 1, but recall accuracy was still high ($M = .78$, $SE = .02$). Updating response time data are shown in Figure 5.

We conducted a $2 \times 2 \times 2$ repeated measures ANOVA with the factors proximity (distant/proximal), digit repetition (no/yes), and CTI (short/long). We found a main effect of proximity, $F(1, 26) = 15.33$, $MSE = .01$, $p < .001$, $\eta_p^2 = 0.37$, and a strong effect of CTI, $F(1, 26) = 105.21$, $MSE = .03$, $p < .001$, $\eta_p^2 = 0.80$. The main effect of digit repetition was not significant, $F(1, 26) = 2.60$, $p = .12$, but the interaction between proximity and digit repetition approached significance, $F(1, 26) = 3.00$, $MSE = .01$, $p < .1$. Inspection of the data in Figure 5 shows that the effect of similarity is essentially a difference between the all-dissimilar (distant/no digit repetition) condition and the three other conditions that each convey some (partial) similarity. In accordance with our a-priori analysis plan, we hence tested the impact of CTI duration on the similarity effect directly by testing the interaction contrast of CTI (short/long) and similarity (the all-dissimilar condition with a lambda weight of 3 vs. the three pooled similar conditions, each with a lambda weight of -1); this interaction contrast was significant, $F(1, 26) = 5.32$, $MSE = .01$, $p = .03$. This demonstrates that the updating RT difference between similar and dissimilar conditions was larger with a short CTI (116 ms) than a long CTI (38 ms).

Discussion

The similarity effect found with a short CTI replicates the findings of Lendinez et al. (2011). We assume that similarity effects in memory updating arise because of representational

overlap between the replaced and the new item. That is, two similar numbers share a digit and/or a region in number space, and to the degree that only new features are substituted, not entire item representations, this similarity will facilitate updating. One could assume that removal of the old item and encoding of the new stimulus occur (in part) in parallel, such that encoding can build on the shared features of the (not yet completely) removed old item. Yet, the more an item representation is removed before the updated number can be encoded, the less facilitation there will be. In support of this, Experiment 2 demonstrated that a well-documented similarity effect in memory updating is diminished when participants are given time to remove a to-be-updated item from memory before encoding the new item. This supports our notion of active removal.

Further support comes from the finding that updating RTs were again substantially reduced with a long CTI (around 280 ms for low similarity updates and 200 ms for high similarity updates). We note that in contrast to the item-repetition condition of Experiment 1, the high-similarity conditions of Experiment 2 did require the substitution of information in memory, that is, both removal of the old and encoding of the new item. Hence the updating RT reduction with a long CTI was found for both high- and low-similarity updates. We argue that this data pattern also rules out the alternative explanation that the reduction of RT with long CTI results from an increased preparedness to update, because preparedness could not explain the strong reduction of the similarity effect found in Experiment 2.

Having established some support for our notion of removal, we turned to the question whether the time it takes to remove a piece of information from WM—that is, removal speed—can be employed as an individual differences measure to test how removal of outdated information from WM might correlate with other variables. Specifically, we argue that the speed-up in updating with vs. without the opportunity to remove ahead of encoding provides a relatively

pure measure of an individual's removal speed, which we can then use to investigate how WM removal abilities co-vary with general WM capacity.

Experiment 3

In Experiment 3, we demonstrate that our removal task can be used to calculate a reliable estimate of removal speed, and we consider this a measure of WM updating ability. The aim of Experiment 3 was to investigate individual differences in removal speed. As reviewed in the Introduction, previous research has emphasized the importance of WM updating as a predictor of higher cognitive functions. This view might suggest a correlation between a WM updating measure and WM capacity, in particular if WM capacity is considered “executive attention” (Kane, Conway, Hambrick, & Engle, 2007).

In contrast, Ecker et al. (2010) identified a substitution component of WM updating that did not co-vary with WM capacity (see Introduction and Figure 1). As we have argued in Ecker et al., removal is a major component of the substitution process. Our prediction was hence that removal speed and WM capacity would be unrelated. Experiment 3 used the removal paradigm from the first two studies (without manipulating item repetition or similarity) and additionally measured WM capacity in a large-scale individual-differences study.

Methods

Participants. Participants ($N = 176$) were undergraduates from the University of Western Australia. None had participated in Experiment 1 or 2. Based on various outlier criteria (see below), nine participants were excluded from the analyses, yielding a final sample of $N = 167$ (125 females, 42 males, mean age 21.8 years, age range 18-55).²

Apparatus, Stimuli, Design, and Procedure. Participants performed an updating task similar to the one described in Experiment 1. The only major modification was that there were no

item repetitions; hence the sole factor of the experiment was CTI (200 vs. 1500 ms). There were 28 trials in total (plus 2 practice trials), with an average of 9 updating steps per trial, yielding approximately 252 updating steps, or 126 per condition. Each trial took approximately 40 seconds, and the experiment took about 20 minutes.

Additionally, participants completed a WM capacity task battery (Lewandowsky, Oberauer, Yang, & Ecker, 2010). This task battery comprises four WM tasks: an operation span task (OS), a sentence span task (SS), a memory updating task (MU), and a spatial short-term memory task (SSTM). We describe the tasks only briefly because we used the default settings described in detail in Lewandowsky et al. (2010).

The two complex-span tasks arguably are the most widely used tasks to measure WM capacity (Conway et al., 2005). They require participants to memorize a set of items for serial recall, while interleaving the memoranda with a secondary processing task such as judging the correctness of equations (OS) or sentences (SS). The two complex-span tasks used memory set sizes of 4 to 8 consonants, and a single arithmetic equation with operands ranging from 1-10 (OS) or a short sentence (SS) in between each pair of memoranda. The MU task was a “standard” WM updating task—measuring mainly generic WM abilities as discussed in the Introduction. It involved encoding a set of between 3 and 5 digits, presented in individual frames on the screen, which were then repeatedly updated before a final cued recall. Updating was prompted by a series of between 2 and 6 successive cues for arithmetic operations (ranging from -7 to +7) presented randomly in individual frames. The final task (SSTM) required participants to memorize and then reproduce a spatial pattern of between 2 and 6 dots, presented sequentially in a 10 × 10 grid. Scores on these 4 tasks can be combined into a single WM score, which has been shown to be a

valid and reliable estimate of WM capacity (Craig & Lewandowsky, 2012; Lewandowsky et al., 2010; Little, Lewandowsky, & Craig, 2012; Sewell & Lewandowsky, 2012).

Results

We excluded participants scoring more than 3 standard deviations from the mean on any of the WM capacity tasks or the recall portion of the updating task, as well as one participant failing to follow instructions, from all analyses.

Updating task. As in Experiment 1, recall accuracy was very high ($M = 0.95$; $SE = 0.003$). A one-way repeated measures ANOVA on updating RTs yielded a significant main effect of CTI, $F(1, 166) = 367.14$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = 0.69$. Replicating the data from Experiment 1, updating took significantly longer with a short CTI ($M = 1018$ ms; $SE = 26$ ms) as compared to a long CTI ($M = 841$ ms; $SE = 24$ ms). This basically replicates Experiments 1 and 2 in that there is a substantial updating RT reduction (of about 180 ms) with a long CTI.

At first glance, this difference between the means of the short and long CTI conditions may seem like an obvious measure of removal speed. However, to calculate an estimate of removal speed that is not confounded with general processing speed (which itself tends to correlate with WM capacity; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000; Wilhelm & Oberauer, 2006, but see Conway, Cowan, Bunting, Theriault, & Minkoff, 2002), we calculated a first removal measure as a proportional gain score³:

$$\text{Removal speed} = (\text{mean}(\text{short CTI}) - \text{mean}(\text{long CTI})) / \text{mean}(\text{short CTI}).$$

To make sure we could accurately estimate the reliability of this score, and also to allow the application of structural equation modeling (see below), we based calculations on random thirds of the data. The mean proportional gain score (obtained by averaging across the three means for each third of trials) was $M = 0.17$ ($SE = 0.007$; Range = -0.16 – 0.40). In a second step,

we tested whether this removal score was statistically reliable. The score's split-third reliability (after Spearman-Brown correction) was $\rho = .71$ and hence sufficiently high to permit further analysis of this score's relation with WM capacity.

The proportional gain score has two limitations, though: (1) It is a rather conservative estimate, in that with a constant RT difference between long and short CTI conditions it decreases as overall RT increases, which could reduce the chance of finding a correlation with WM capacity; and (2) its reliability might in part be driven by the reliability of overall RT. Hence we calculated an alternative removal score, namely the individual residuals from a simple linear regression model predicting the short CTI RTs from the long CTI RTs (cf. Oberauer, Lange, & Engle, 2004). Again, this was done using random thirds of the data. This score is zero for a person with average removal speed and grows positive (negative) for a person with removal speed higher (lower) than predicted from their overall updating speed. Thus, the mean regression-residual score was $M = 0.00$ s ($SE = 0.009$ s; Range = -0.52 s – 0.40 s). This score's split-third reliability (after Spearman-Brown correction) was $\rho = .77$ and hence also sufficiently high to permit further analysis in relation to WM capacity.

WM capacity battery. Table 1 gives descriptive data of the WM capacity tasks. From the four primary task scores, we calculated a global WM capacity score for each participant⁴ (cf. Lewandowsky et al., 2010).

Correlational analyses. As expected, all four WM capacity scores correlated significantly with each other, with r ranging from 0.26 (OS and SSTM) to 0.76 (OS and SS). The global WM capacity score also correlated substantially with recall accuracy in the updating task ($r = .38$, $p < .001$). Both short-CTI and long-CTI RTs had a relatively low correlation with WM capacity ($r = -.16$, $p = .04$ for short CTI and $r = -.14$, $p = .07$ for long CTI).

We then calculated the correlation between the global WM capacity score and the two removal scores, viz. the proportional gain score and the regression residual score. We additionally excluded one bivariate outlier (Mahalanobis distance probability < 0.001) from the analysis of the regression-residual score (although this did not affect the outcome significantly). Using the proportional gain score as an index of removal speed, the correlation between WM capacity and removal ability was virtually zero, $r = -.02$; $p = .84$; 95% $CI = -.17 - .14$. Using the regression residual score as an index of removal speed, the correlation between WM capacity and removal ability was likewise negligible, $r = -.07$; $p = 0.39$; 95% $CI = -.22 - .08$. Moreover, neither of the removal scores correlated significantly with any of the individual WM tasks (the correlations with the OS, SS, MU, and SSTM tasks were $-.04$, $-.03$, $.04$, and $-.07$ for the proportional gain score, and $-.07$, $-.06$, $-.04$, and $-.06$ for the regression-residual-based score, respectively, all p 's > .35). Scatterplots plotting WM capacity and the proportional-gain and regression-residual-based removal speed indices, respectively, are presented in Figure 6 and 7.

To further corroborate this non-significant relation between WM capacity and removal efficiency, we employed structural equation modeling (see Figure 8). We calculated a latent WM capacity (WMC) factor from the four WM capacity tasks (OS, SS, MU, SSTM), and a latent estimate of removal efficiency (REM) from randomly selected thirds of the proportional-gain removal scores (PG1 to PG3; upper panel of Figure 8) and the regression-residual-based removal scores (RR1 to RR3; lower panel of Figure 8), and linked the two latent factors. The model using the proportional gain scores achieved a very good fit to the data, $\chi^2(12) = 12.57$; $RMSEA = .017$; $CFI = .998$. The correlation between the latent WM capacity and removal factors was non-significant, $r = -.04$, $p = .67$. The model using the regression-residual-based score achieved an

excellent fit to the data, $\chi^2(12) = 9.96$; $RMSEA = 0$; $CFI = 1$. The correlation between the latent WM capacity and removal factors was non-significant, $r = -.07$, $p = .44$.

Discussion

The aim of Experiment 3 was to investigate the relation between removal speed and WM capacity. WM capacity has been termed the “engine of cognition” (Jonides, 1995); it accounts for roughly half the variance in general fluid intelligence (Oberauer et al., 2005). Previous research has demonstrated that WM updating task performance is strongly related to performance on WM capacity tasks (Schmiedek et al., 2009). We have argued that the reason for this is that “standard” WM updating tasks mainly measure WM capacity, whereas only the removal process identified in the present experiments is unique to memory updating. In this vein, we argue that our removal scores did not even correlate with performance on the MU task because the MU score is an accuracy score driven mainly by generic WM processes such as retrieval and transformation (cf. Ecker et al., 2010). In other words, while removal efficiency is likely to be an additional factor determining performance on the MU task, it is only a minor contributor, and slow removers may not necessarily perform (much) less accurately on the MU task. In sum, the results of our individual differences study indicate that there is no relation between WM capacity and removal speed, which is a more specific measure of WM updating.

General Discussion

In this article we have introduced a novel measure of WM updating. Traditional WM updating tasks arguably measure general WM processes in addition to updating, whereas it is the removal of information from WM that is specific and unique to WM updating. We demonstrated that giving people preparation time to remove information from WM speeds up updating when new information is subsequently presented. Our notion of removal by unlearning item-position

associations is a specific incarnation of the more general idea advanced by Kessler and Meiran (2008), who suggested that partial updating of memory sets involves “dismantling” or “unbinding” the old representations.

Whereas our research is guided by the SOB model, and our data support the implementation of removal into SOB (Oberauer et al., 2012), it is important to note that other researchers have provided independent evidence for the existence of an active and attention-demanding removal process. For example, Fawcett and Taylor (2008, 2012) have shown that directed forgetting of an item (1) slows down responses on an unrelated secondary task for up to 2.6 seconds after the forget cue, and (2) impairs incidental memory for a subsequent distractor, in particular when the directed forgetting of the studied item is successful.

In addition to the experimental evidence for removal of outdated material from WM, we found that the speed of removal can be reliably measured, and that it does not covary with WM capacity. The independence of removal speed from WM capacity confirms and specifies the earlier finding that the ability to substitute information in WM is unrelated to WM capacity (Ecker et al., 2010). This is a noteworthy finding because WM capacity has been found to predict a broad range of cognitive abilities from categorization (Craig & Lewandowsky, 2012) to metaphor comprehension (Chiappe & Chiappe, 2007). The removal measure is therefore a highly specific index of WM updating ability.

This result further suggests that the previously reported predictive link between WM updating and fluid intelligence (Chen & Li, 2007; Friedman et al., 2006) can be explained by the fact that the utilized WM updating tasks required generic WM processes that were unrelated to the actual updating process. Further research is needed to ascertain whether removal in itself predicts some intelligence-related variance (cf. Colzato, van Wouwe, Lavender, & Hommel,

2006), or whether it relates to other executive functions (e.g., inhibition or filtering of information) or real-world updating on longer time-scales (e.g., Ecker, Lewandowsky, Swire, & Chang, 2011).

We also note that the present findings contradict an assumption made by Oberauer et al. (2012) in their simulation of correlations between span tasks in the framework of the SOB-CS model. They argued that removal rate is the source of the common variance across complex-span tasks with different content domains. While it is likely that removal ability is an important determinant of complex-span performance—because a complex-span task will require the removal of distractors during encoding and removal of already-recalled list items during recall—the present lack of correlation between removal ability and WM capacity suggests that there are other sources of variance that could explain the correlation between span tasks.

These statements should not, however, imply that removal is unimportant. On the contrary, we assume removal to be crucial to maintaining a functional WM system that can efficiently focus on relevant information. Theoretically, it is vital for a WM system to have a ‘housekeeping’ mechanism in place that allows no longer relevant information to leave the system in order to prevent the build-up of interference. In many models this mechanism is proposed to be passive and constant decay (Baddeley, 2000; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). Yet, in line with recent research that has failed to find support for the existence or relevance of decay in WM (Berman et al., 2009; Jalbert et al., 2011; Lewandowsky et al., 2009; Oberauer & Lewandowsky, 2008, 2013; see also Ecker & Lewandowsky, 2012), we propose that a flexible and active removal process is an ideal candidate to fill the role of this ‘housekeeping’ mechanism in WM.

References

- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 570-585.
- Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 35, 317-333.
- Brainard, D. H. (1997). The Psychophysics toolbox. *Spatial Vision*, 10, 433-436.
- Chen, T., & Li, D. (2007). The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Aging, Neuropsychology, and Cognition*, 14, 631-646.
- Chiappe, D. L., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, 56, 172-188.
- Chuderski, A., Taraday, M., Nęcka, E., Smoleń, T. (2012). Storage capacity explains fluid intelligence but executive control does not. *Intelligence*, 40, 278-295.
- Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, 36, 584-606.
- Colzato, L. S., van Wouwe, N. C., Lavender, T. J., & Hommel, B. (2006). Intelligence and cognitive flexibility: Fluid intelligence correlates with feature “unbinding” across perception and action. *Psychonomic Bulletin & Review*, 13, 1043-1048.

- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163-183.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769-786.
- Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *Quarterly Journal of Experimental Psychology, 65*, 439-464.
- De Beni, R., & Palladino, P. (2004). Decline in working memory updating through ageing: Intrusion error analyses. *Memory, 12*, 75-89.
- Ecker, U. K. H., & Lewandowsky, S. (2012). Computational constraints in cognitive theories of forgetting. *Frontiers in Psychology, 3*, 400.
- Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2013). Removal of information from working memory. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 400-405). Austin, TX: Cognitive Science Society.
- Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 36*, 170-189.
- Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review, 18*, 570-578.

- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309-331.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*, 59-79.
- Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, *36*, 1168-1181.
- Fawcett, J. M., & Taylor, T. L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica*, *139*, 84-90.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172-179.
- Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition*, *26*, 263-276.
- Jalbert, A., Neath, I., Bireta, T. J., & Surprenant, A. M. (2011). When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 338-353.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*, 138-202.
- Jonides, J. (1995). Working memory and thinking. In E. E. Smith & D. N. Osherson (Eds.), *Invitation to cognitive science: Thinking* (Vol. 3). Cambridge, MA: MIT Press.

- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21-48). Oxford: Oxford University Press.
- Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 1339-1348.
- Kessler, Y., & Oberauer, K. (2013). *Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation*. Manuscript submitted for publication.
- LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience*, *25*, 127-142.
- Lendinez, C., Pelegrina, S., & Lechuga, T. (2011). The distance effect in numerical memory-updating tasks. *Memory & Cognition*, *39*, 675-685.
- Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 49, pp. 1-48). London: Elsevier.
- Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, *13*, 120-126.
- Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. H. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, *42*, 571-585.

- Little, D. R., Lewandowsky, S., & Craig, S. (2012). *Working memory capacity and fluid abilities: The more difficult the item, the more more is better*. Manuscript submitted for publication.
- Martínez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, M. T., & Colom, R. (2011). Can fluid intelligence be reduced to 'simple' short-term storage? *Intelligence, 39*, 473-480.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, T. D., & Wager, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*, 61-64.
- Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 948-957.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 411-421.
- Oberauer, K., Lange, E., & Engle, R. W. (2004). Working memory capacity and resistance to interference. *Journal of Memory and Language, 51*, 80-96.
- Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 544-576.
- Oberauer, K., & Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General, 142*, 380-411.

- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*, 779-819.
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 61-65.
- Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity—facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017-1045.
- Schmiedek, F., Hildebrandt, A., Lovdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 35*, 1089-1096.
- Sewell, D. K., & Lewandowsky, S. (2012). Attention and working memory capacity: Insights from blocking, highlighting, and knowledge restructuring. *Journal of Experimental Psychology: General, 141*, 444-469.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1436-1451.
- Wilhelm, O., & Oberauer, K. (2006). Why are reasoning ability and working memory capacity related to mental speed? An investigation of stimulus-response compatibility in choice reaction time tasks. *European Journal of Cognitive Psychology, 18*, 18-50.
- Yntema, D. B., & Mueser, G. E. (1962). Keeping track of variables that have few or many states. *Journal of Experimental Psychology, 63*, 391-395.

Footnotes

¹ We acknowledge the fact that repetition of the first position occurred only in proximal updates, while repetition of the second position occurred only in distant updates. Whether digit repetition effects depend on the repetition's position may be an interesting topic for future research but was beyond the scope of the present article.

² Removing age outliers did not affect the outcome of the analyses reported below.

³ Using the long-CTI condition as the baseline yielded the exact same result pattern.

⁴ We retained the MU task in this global score [unlike Ecker et al. (2010)] despite its inherent similarity to the present updating task. We note that this works against our hypothesis that removal ability (estimated from our present updating task) should not correlate with WM capacity. Excluding MU task performance from the global WM capacity score did not affect the outcome of the analyses reported below.

Table 1

Descriptive Statistics of Performance on the Working Memory Tasks in Experiment 3

Task	Mean	<i>SE</i>	Range	Skewness	Kurtosis
OS	0.72	0.009	0.33 – 0.96	-0.55	0.44
OS _{pt}	0.92	0.005	0.61 – 1.00	-1.65	3.56
SS	0.69	0.010	0.31 – 0.97	-0.60	0.17
SS _{pt}	0.92	0.003	0.78 – 1.00	-0.88	0.94
MU	0.65	0.014	0.17 – 0.97	-0.41	-0.48
SSTM	0.85	0.004	0.69 – 0.98	-0.38	0.15

Legend. OS, Operation Span; SS, Sentence Span; *pt* denotes processing tasks; MU, Memory Updating; SSTM, Spatial Short-Term Memory; *SE*, Standard error of the mean; $N = 167$.

Figure Captions

Figure 1. Graphical representation of the structural equation model of updating accuracy data from Ecker et al. (2010), showing the prediction of latent updating factors GenAcc (general accuracy), R (retrieval), T (transformation), and S (substitution) by a latent working memory capacity (WMC) factor. Manifest accuracy variables reflect log-transformed accuracy data referring to Ecker et al.'s eight experimental conditions, with bold, capital letters implying the process was involved in the condition, and small letters with a 'no' subscript indicating the process was not involved in the condition (e.g., the experimental condition involving all three processes is labeled **R-T-S**; the condition featuring only a substitution is labeled *r_{no}-t_{no}-S*, etc.). The WMC-related manifest variables reflect mean performance in WM capacity tasks OS (operation span), SS (sentence span), and SSTM (spatial short-term memory). Estimated standardized weights (correlations, in boldface) are presented adjacent to latent connections. Estimated unstandardized means (in log-accuracy units, italicized) are shown inside the latent factors. Means of latent factors that are not given in the figure (error variables and WMC factor) were fixed at 0. Regression weights in the working memory updating (WMU) measurement model were fixed at 1, with the exception of the link between T and the **R-T-s_{no}** variable, which was freely estimated (dashed arrow with unstandardized estimate in italics). All estimated covariances provided in the figure are (marginally) significant, $p < .051$; all estimated means are significantly different from 0, $p < .001$. e1 – e11 = error variables.

Figure 2. A trial sequence from Experiment 1. The example trial features 3 updating steps. Across trials, the number of updating steps ranged from 1 to 21, with a 10% termination probability after each step; this yielded a mean number of approximately 9 updating steps per

trial. Note that the duration of the empty-frames response-cue interval at the beginning of each updating step was determined by the length of the subsequent CTI: If the CTI was short (200 ms), the response-cue interval was long (1800 ms), if the CTI was long (1500 ms), the response-cue interval was short (500 ms), and hence the retention interval between updating steps was constant at 2000 ms.

Figure 3. A schematic depiction of an updating step in the long cue-target interval (CTI) condition (top panel) and the short CTI condition (bottom panel) of Experiment 1. After a blank-frames screen in the response-cue interval, the removal cue (bold red frame) was given during the CTI, before the to-be-encoded letter was presented until a response was given. The duration of the response-cue interval was determined by the subsequent CTI.

Figure 4. Updating response times from Experiment 1. Vertical bars denote within-subject standard errors of the mean (Morey, 2008).

Figure 5. Updating response times from Experiment 2. Vertical bars denote within-subject standard errors of the mean (Morey, 2008).

Figure 6. The correlation between a working memory capacity score (scale 0 to 1) and a removal speed index. The removal speed index is based on proportional gain scores and can be interpreted as the increase in RT associated with the requirement to remove an item during updating, relative to overall updating RT (in seconds; see text for details).

Figure 7. The correlation between a working memory capacity score (scale 0 to 1) and a removal speed index. The removal speed index can be interpreted as a deviation from the average removal speed (in seconds), with negative values indicating higher removal speed (see text for details).

Figure 8. Graphical representation of the two structural equation models applied to the data of Experiment 3, showing the non-significant relation between WM capacity (WMC) and removal efficiency (REM). Manifest variables related to WMC reflect mean performance in WM tasks OS (operation span), SS (sentence span), MU (memory updating), and SSTM (spatial short-term memory). Manifest variables related to REM reflect proportional-gain-based removal speed indices (upper panel) and regression-residual-based removal speed indices (lower panel) derived from randomly selected thirds of the updating task data of Experiment 3. Values adjacent to unidirectional connections reflect estimated standardized regression weights; values adjacent to bidirectional connections reflect correlations; e1-e7 = error variables.

Figure 1

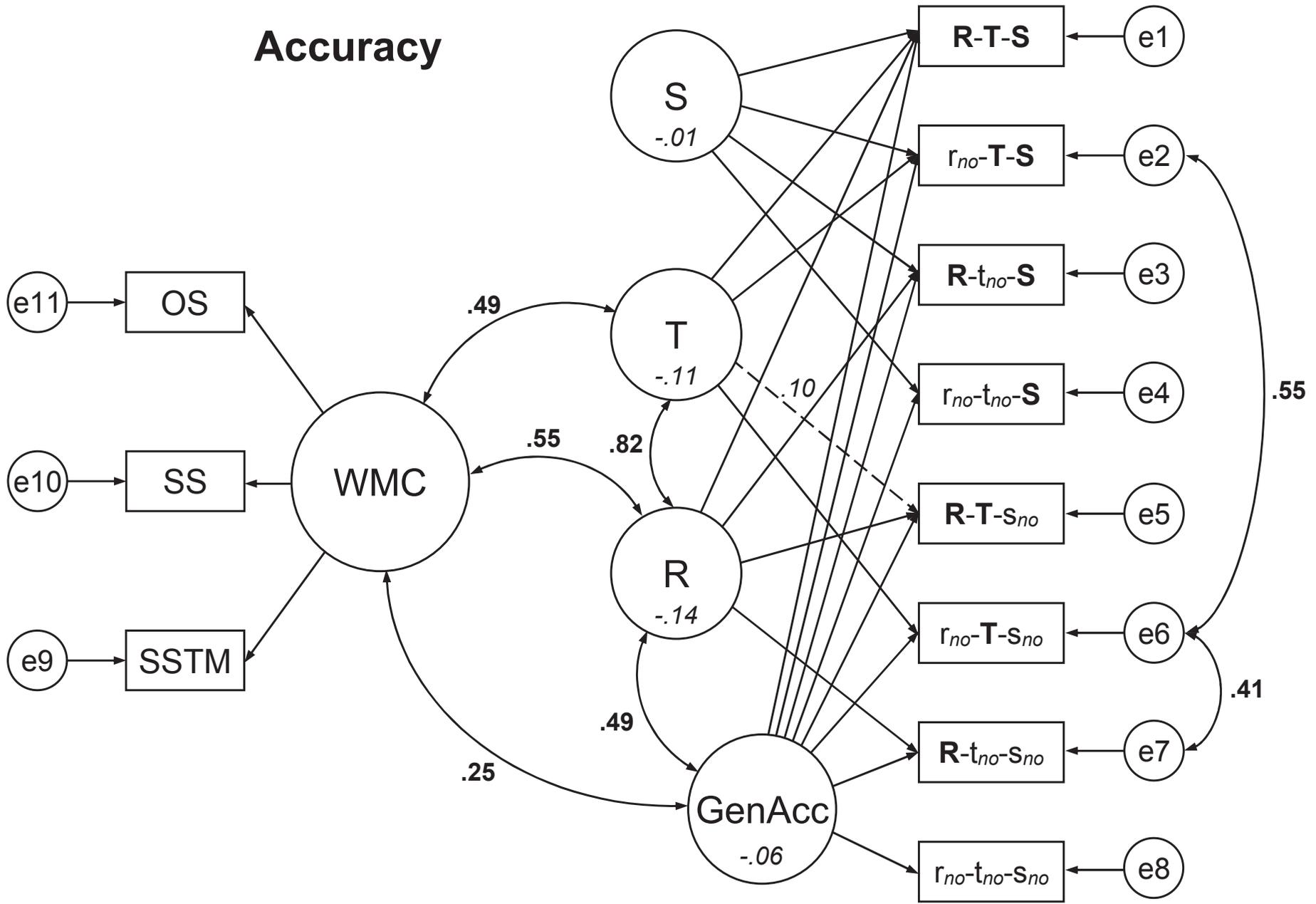


Figure 2

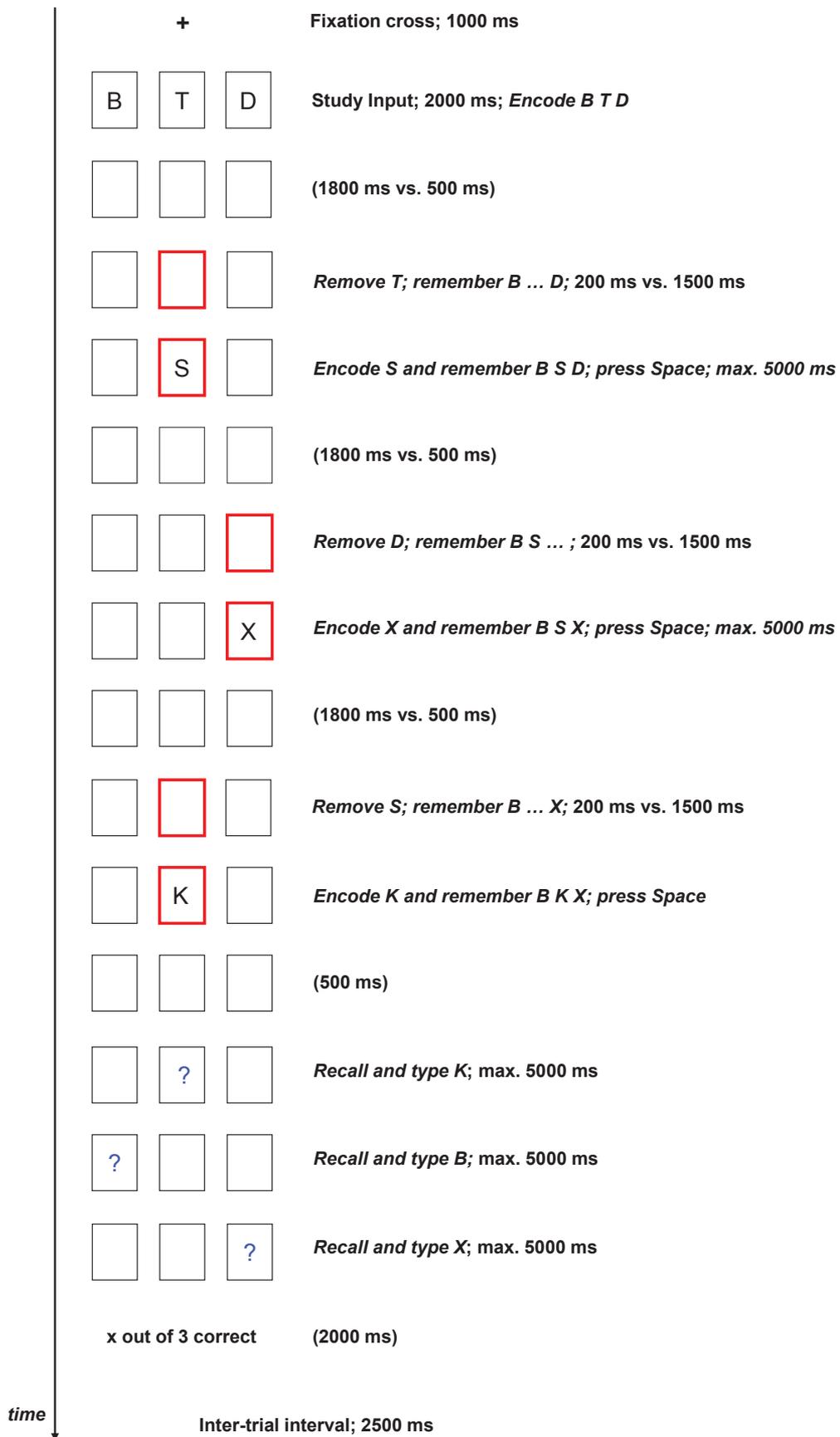


Figure 3

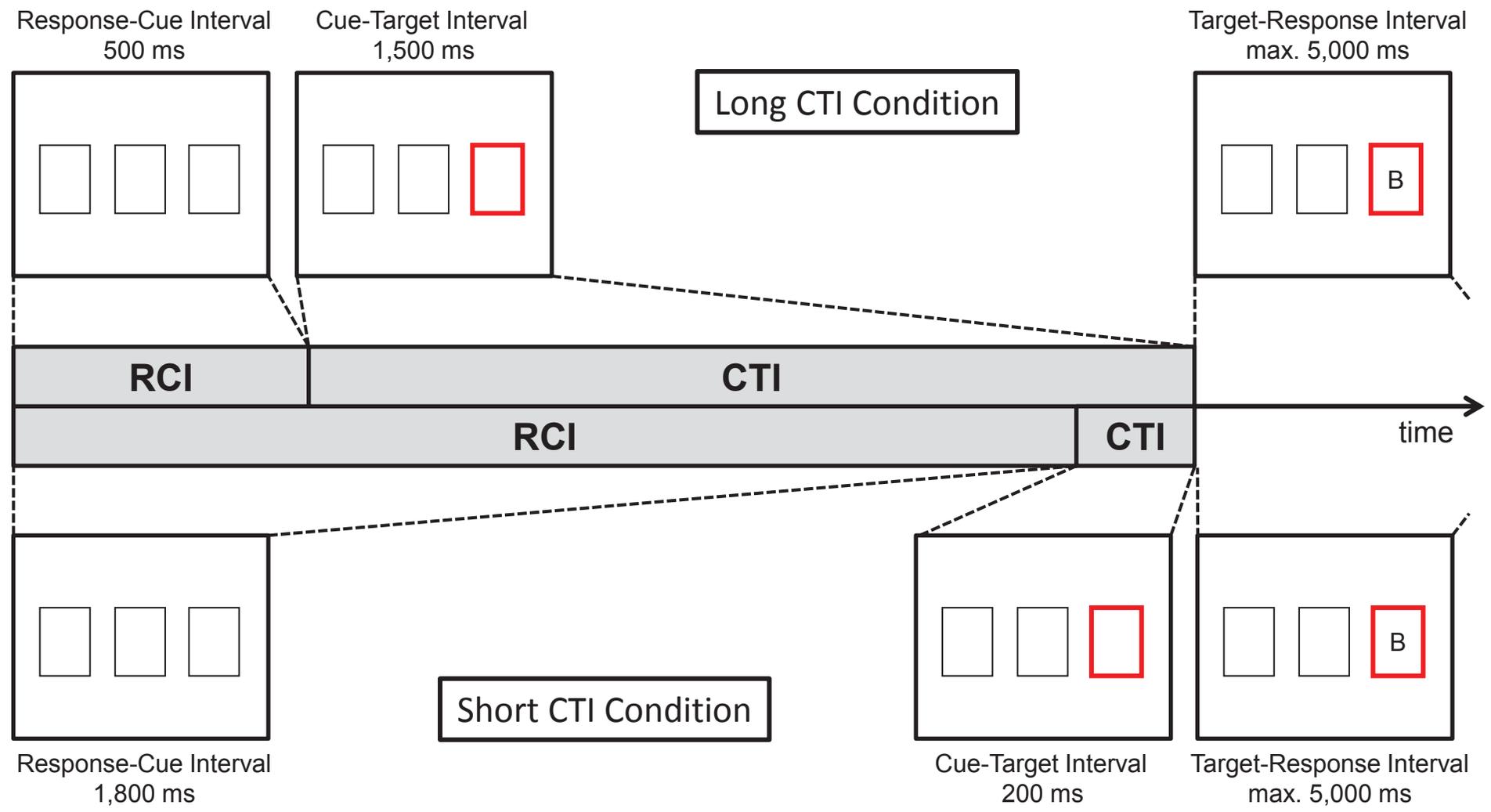


Figure 4

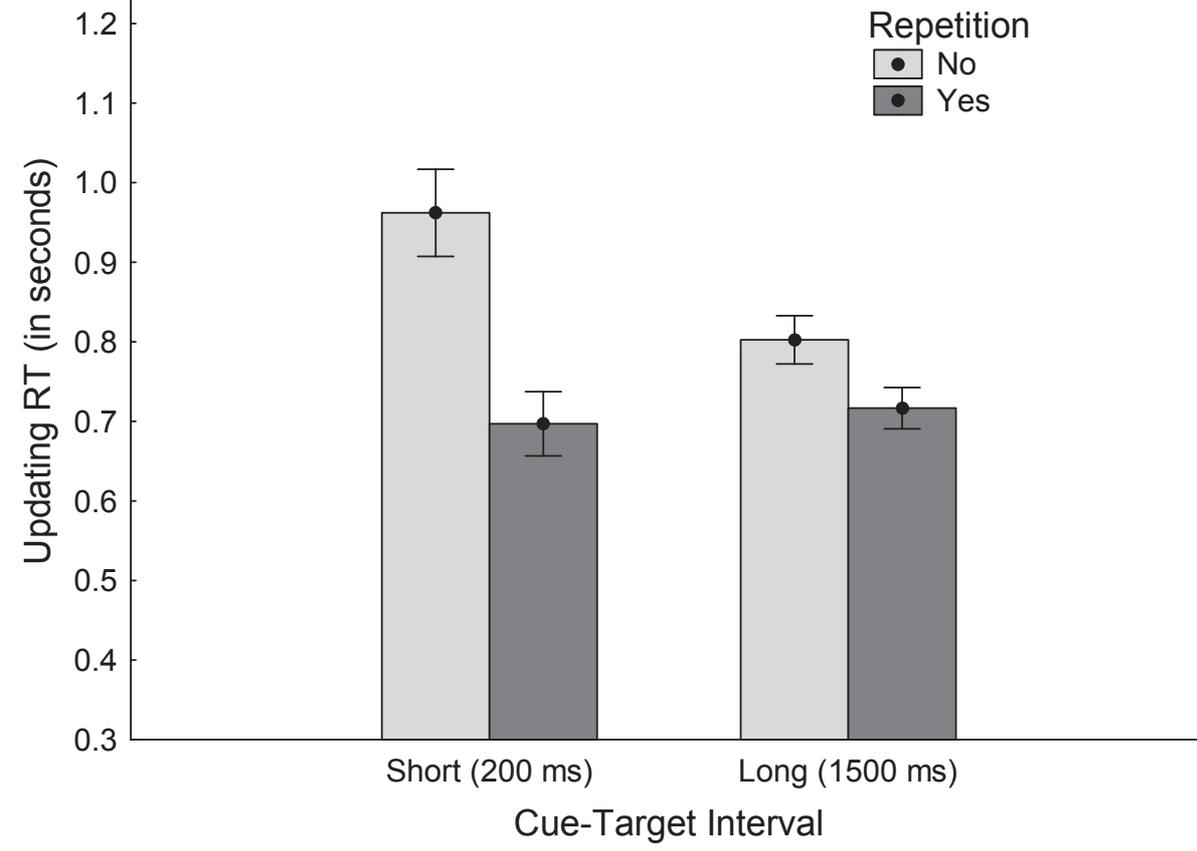


Figure 5

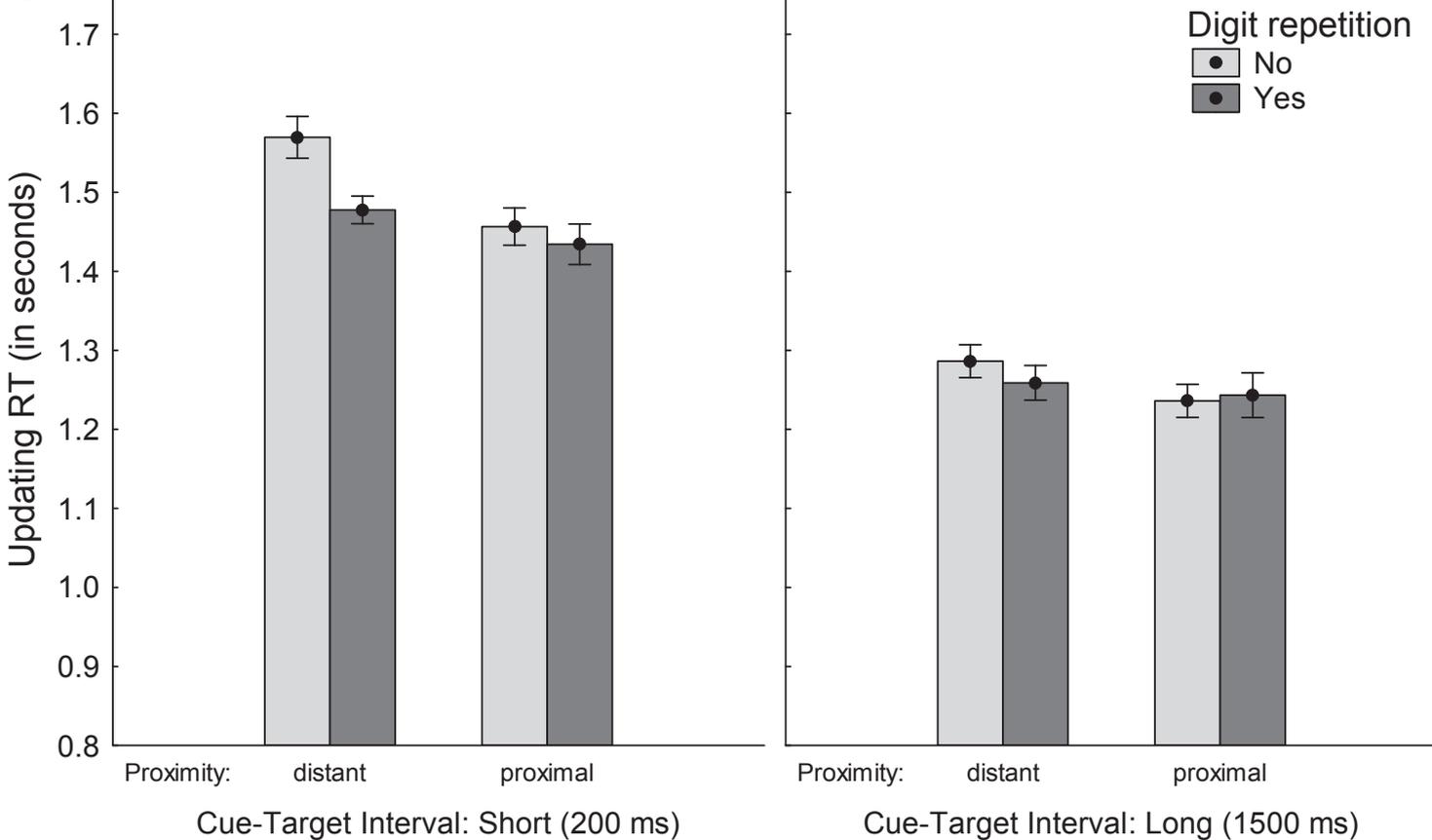


Figure 6

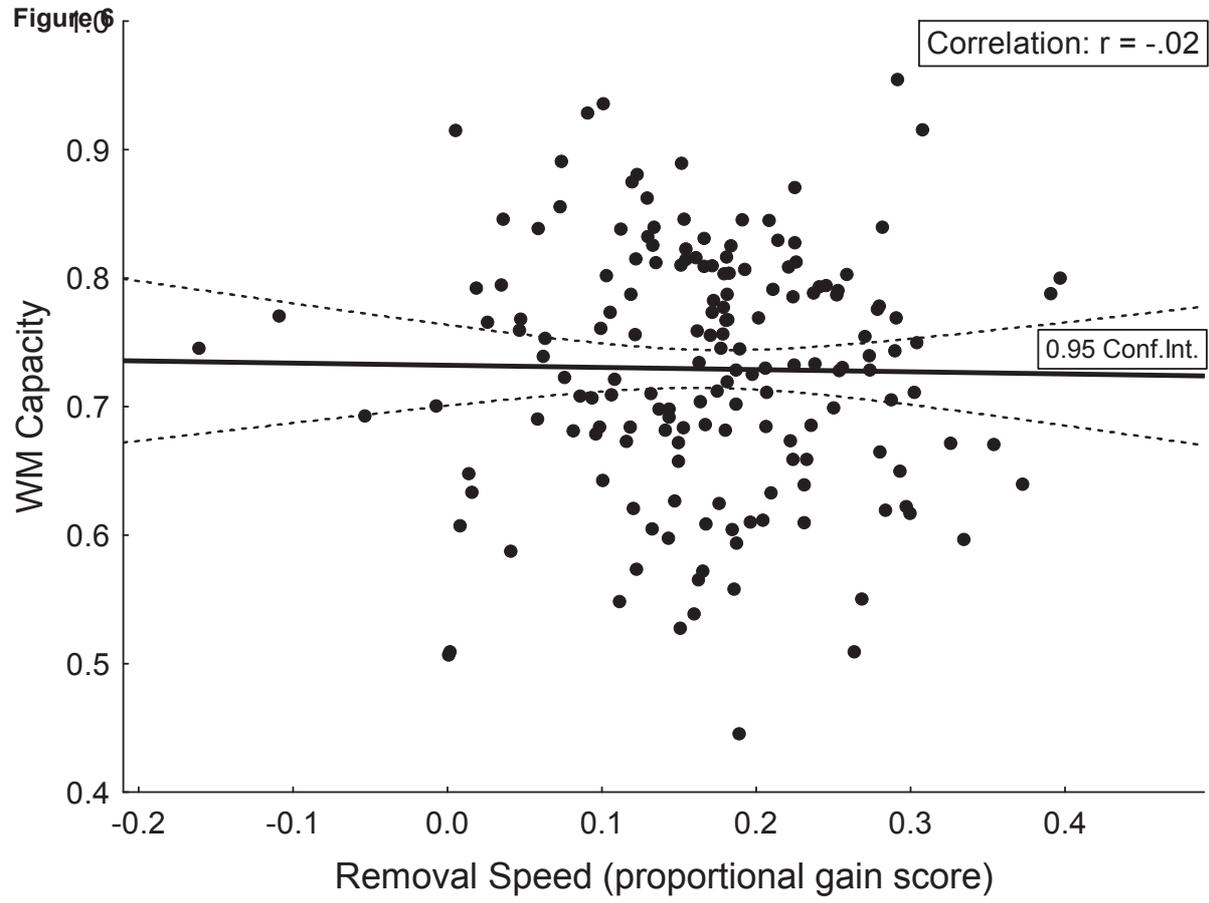


Figure 7

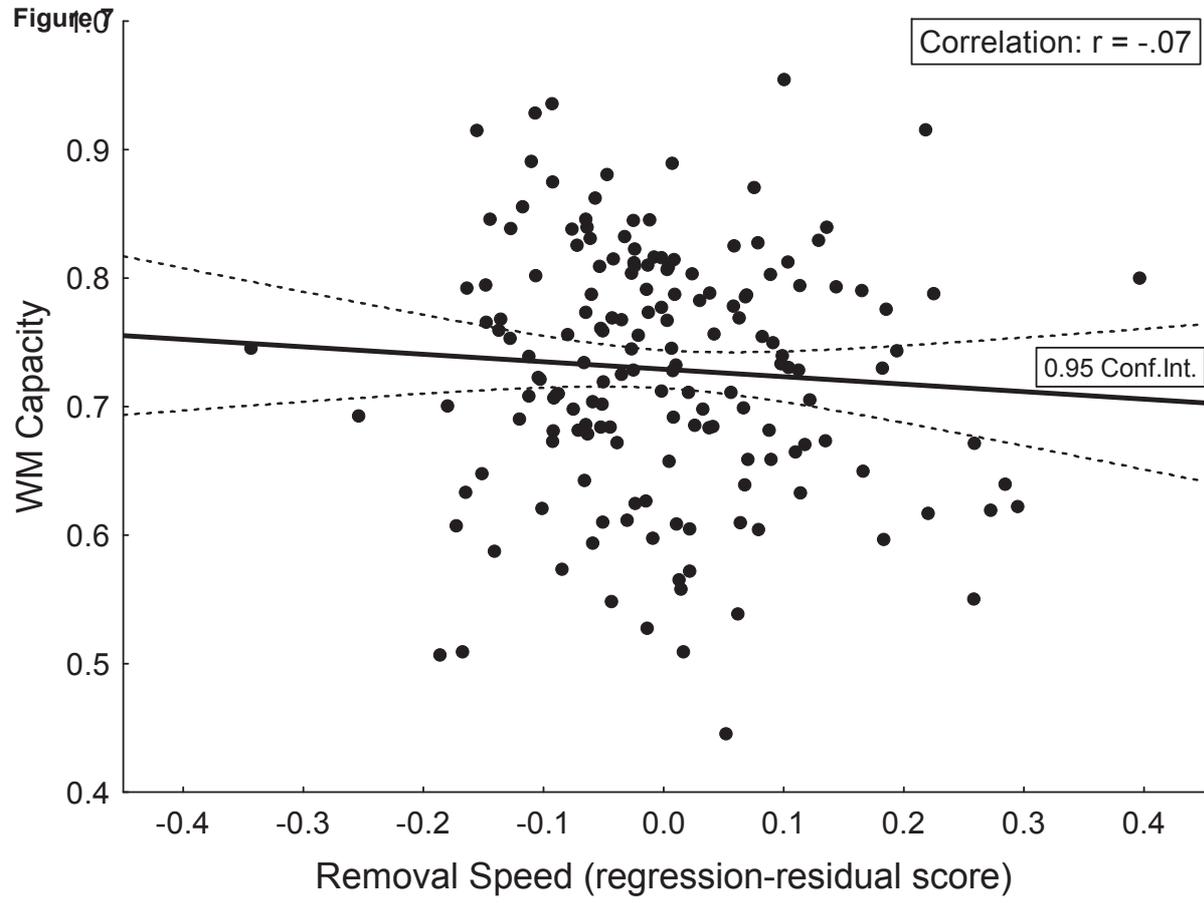


Figure 8

