The slow rate of visual working memory consolidation is a structural limit

by

Brandon Carlos

A thesis submitted to the Department of Psychology,

College of Liberal Arts and Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

Chair of Committee: Benjamin Tamber-Rosenau

Committee Member: Paul Cirino

Committee Member: Hanako Yoshida

University of Houston

May 2020

ABSTRACT

Extensive research has focused on the limited storage capacity of working memory (WM), i.e., the maximum amount of information that can be maintained in WM. However, a relatively understudied limitation of WM involves the processing speed by which sensory information can be transformed into a WM representation that is resistant to distraction from ongoing perception and cognition. The speed of this "consolidation" process is the subject of conflicting results. Researchers have arrived at estimates of the consolidation time course using distinct paradigms ranging from 25 ms to 1 s, meaning more than an order of magnitude of variability. The extremely large variation in WM consolidation speed estimates across measurement approaches motivated the current work's goal of determining whether consolidation speed is under strategic control or is a stable structural constraint of WM encoding. Here, the slower (1 s) measurement of WM consolidation of visually-presented verbal stimuli (i.e., letters) was replicated by using retroactive interference (RI; Nieuwenstein & Wyble, 2014)—essentially, measuring how long it takes after a WM sample array is presented for the representation in WM to no longer be vulnerable to distraction by performing a speeded second task (T2). Then, the RI results were extended to more standard visual WM stimuli (i.e., color patches). Further, slow consolidation was obtained regardless of the relative prioritization of WM encoding vs. T2, supporting the structural account. However, no RI was obtained when T2 was unspeeded. Finally, a sensorimotor decision and motor response to T2 were required to obtain RI. Given that RI was robust to varying WM probes, WM stimuli, and that slow consolidation was obtained regardless of strategic demands, the present study supports the structural account of WM consolidation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The mechanisms and limitations of human working memory (WM)—the mental workspace that allows us to maintain and manipulate sensory and other goal-relevant information—are hotly debated. Though recent debates center on the quantity and precision of information that WM stores, another unsolved dilemma is the temporal dynamics of WM consolidation, i.e., the time it takes for information to be stored into WM. For instance, some studies have used visual masking to show that consolidation is rapid, taking only about 20-50 ms per item (Gegenfurtner & Sperling, 1993; Vogel, Woodman, & Luck, 2006; Woodman & Vogel, 2005). However, psychological refractory period (PRP)/dual-task proactive interference (PI) experiments suggest the duration of consolidation is around 300 ms (Gegenfurtner & Sperling, 1993; Jolicoeur & Dell'Acqua, 1998). Finally, attentional blink studies have shown that consolidation of a single letter can take hundreds of milliseconds, suggesting a much slower consolidation process (Chun & Potter, 1995; Raymond, Shapiro, & Arnell, 1992). Thus, estimates of consolidation time vary across measurement approaches by an order of magnitude.

The process by which information is stored into WM begins with incoming sensory input. In the case of visual input, all of the visual information in one's visual field is temporarily registered via iconic memory, which is a type of sensory memory system that has an unlimited capacity, but is very short lived (Gegenfurtner & Sperling, 1993; Neisser, 1967). Because the high capacity cannot be sustained for long, humans must determine what to direct our attention to in order to consolidate that subset of the visual information into WM. If successfully consolidated into WM, the information can then endure interference from other incoming visual input for the matter of a few seconds (A.

D. Baddeley & Hitch, 1974; Phillips, 1974). Figure 1 provides an illustration of the process of consolidating information into WM.

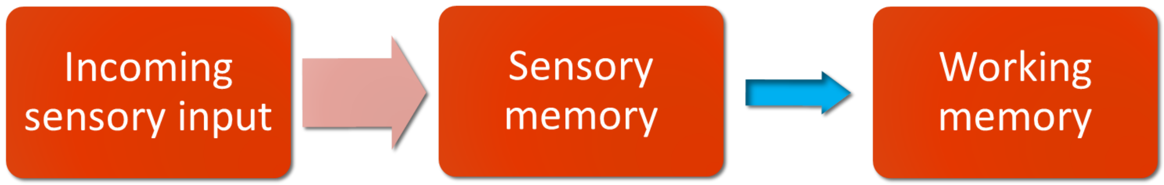

Figure 1. The process of information being stored into WM. The arrow separating the first two boxes (left to right) is big because it represents the unlimited capacity of sensory information that is registered into sensory memory. The blue arrow is smaller because of the limited capacity of WM. The blue arrow also represents the consolidation process.

Due to the many different paradigms used to measure WM consolidation, there have been varying estimates of the time course of consolidation. Rapid estimates come from visual masking paradigms, which rely on the assumption that successful WM is dependent on consolidation having been completed before onset of a mask. This assumption stems from the ideas that consolidation can only continue for as long as the sensory representation (e.g., iconic memory) of an item is available, and that masking diminishes sensory memory (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011; Breitmeyer & Ogmen, 2000; Bundesen, 1990; Shibuya & Bundesen, 1988; Vogel et al., 2006; Zhang & Luck, 2008). By varying the stimulus onset asynchrony (SOA—the duration of time between the presentation of one stimulus to the onset of a second stimulus) between the WM array and the mask, masking paradigms have suggested a rapid consolidation process lasting approximately 50 ms (Vogel et al., 2006; see also Fuller, Luck, McMahon, & Gold, 2005; Woodman & Vogel, 2005).

Alternatively, intermediate estimates of consolidation speed may be derived from the attentional blink (AB), an impairment in the ability to identify the second of two targets separated by approximately 200-500 ms. The AB suggests a longer duration of consolidation (Raymond et al., 1992). The rapid serial visual presentation (RSVP) paradigm is used to study the AB. This paradigm involves having participants identify one or more targets presented in a continuous stream with other distractor stimuli. Studies inferring the temporal dynamics of WM consolidation from the AB assume that if the second target goes undetected, consolidation of the first target must still be ongoing (Lagroix, Spalek, Wyble, Jannati, & Di Lollo, 2012; Shih, 2008; Taatgen, Juvina, Schipper, Borst, & Martens, 2009; Wyble, Bowman, & Nieuwenstein, 2009; Wyble, Potter, Bowman, & Nieuwenstein, 2011). Since there is a significant reduction in accuracy of T2 when presented hundreds of milliseconds after T1, the duration of consolidation must therefore be much longer than what masking paradigms suggest (Bowman & Wyble, 2007; Chun & Potter, 1995; Jolicoeur & Dell'Acqua, 1998). One assumption of such interpretations of the AB is that consolidation must continue after a visual mask, since AB paradigms typically use RSVP in which each stimulus is masked by the next stimulus after approximately 100 ms.

The AB is one example of PI of WM consolidation on another task—in the case of the AB, PI of first target (T1) consolidation on the attentional selection (in time) of a second target (T2) from the RSVP stream and/or on consolidation of T2 into WM. Other dual-task PI approaches eschew the RSVP paradigm and instead present discrete targets at varying time offsets using the psychological refractory period (PRP) approach (Pashler, 1994; Telford, 1931; Welford, 1952). The PRP effect is observed when the response time for the second of two sequential tasks is slowed with decreasing SOAs. With a longer duration of time between T1 and T2, the PRP effect diminishes, returning T2 responses to their typical timing. It is believed that the AB and PRP effect occur from a shared brain mechanism (Marti, Sigman, & Dehaene, 2012). Both phenomena result from the same central processing stage, due to the ongoing processing of T1 using a resource that is also required for the processing of T2. One prominent explanation for the discussed effect regards an immutable central processing bottleneck which prevents the processing of T2 until the processing of T1 has concluded (Marois & Ivanoff, 2005; Pashler, 1994). Thus, by varying SOAs, the time required to process T1 can be inferred by the minimum required SOA for T2 response time to return to its usual speed. It must be noted there are other potential theories which account for the PRP effect. These include resource sharing accounts that assume there is a limited processing capacity that can be shared between tasks, instead of a central bottleneck (Kahneman, 1973; Koch, Poljac, Müller, & Kiesel, 2018; Tombu & Jolicoeur, 2003). While PRP approaches typically examine PI of a sensorimotor decision T1 on a second sensorimotor decision T2, similar results have been obtained with WM encoding as the T1 task (e.g., Jolicoeur & Dell'Acqua, 1998; Tombu et al., 2011).

The slowest estimates of consolidation stem from dual-task retroactive interference (RI) studies. Nieuwenstein and Wyble (2014) sought to resolve the discrepancy between rapid (masking paradigm) and slow (AB paradigm) consolidation estimates by crossing a retroactive dual-task interference manipulation with a masking manipulation. To do so, they examined the interval during which WM consolidation can be disrupted by an intervening task (a speeded two-alternative forced-choice number parity judgment) when the WM array either was or was not immediately followed by a visual mask. They found that a speeded two-alternative forced-choice (2AFC) during the WM delay diminished WM performance, but that this effect abated with increasing WM array-2AFC SOAs to 1 second, suggesting that WM consolidation continued up to 1 second. This is the slowest-yet estimate of the time course of WM consolidation. Their results held true for both letters and complex unfamiliar visual stimuli (Kanji characters), and they obtained the same RI effect regardless of the presence or absence of a mask.

The extremely large variation in WM consolidation speed estimates across paradigms motivated the present question of whether consolidation speed was under strategic control (a result of task instructions, priority, or other mutable factors) rather than being a structural constraint (as implicitly assumed by the comparison of estimates across measurement paradigms). Nieuwenstein and Wyble's results and interpretation supports the structural account (see Figure 2a). The structural account assumes that a second task can interrupt WM consolidation. Manipulating the SOA between T1 and T2 revealed that WM consolidation takes up to 1 second to be complete. This pattern may be explained by the notion that long SOAs allowed full consolidation of the WM sample before T2, while short SOAs led to incomplete consolidation before disruption by T2. A structural account would suggest that a slow process is potentially due to how the WM system is built. However, there is research that supports a potential alternative account (i.e. strategic account; see Figure 2b; cf., Schumacher et al., 2001; Tombu & Jolicoeur, 2003). For the strategic account, if task-specific strategies lead participants to budget some capacity for the expected second task, it would suggest a prolongation of an otherwise rapid process. Knowing a second task is going to happen, instead of using all available resources for

WM processing, only some are used. As a consequence of splitting resources, WM consolidation may proceed slowly, which is why it is vulnerable to interruption up to 1 second. The present study will adjudicate which of these two accounts (structural vs. strategic) are supported by the data that is collected from the original, and slightly manipulated, retroactive interference paradigm.

Furthermore, Nieuwenstein and Wyble's RI results—where performance on the first task is affected by the second task—must be reconciled with prior results that instead showed PI—a PRP effect in which performance on the second task is affected by the first task—in similar dual-task paradigms (e.g., Jolicoeur & Dell'Acqua, 1998; Tombu et al., 2011); this provides further motivation to study factors mediating between PI and RI in dual-task WM encoding paradigms. It must be noted that the conventional PRP-style experiments differ from RI-style experiments in that consolidation rate is inferred from PI of encoding a sample array into memory on a subsequent T2 decision task. It can be argued that this measure provides indirect evidence about the rate of consolidation because it assumes that it is the consolidation process itself, rather than ancillary processes related to WM, that drive interference. On the other hand, Nieuwenstein and Wyble's RI-style studies rely on a weaker assumption that something about an intervening decision task is interfering with WM consolidation. However, it is not clear if structural or strategic factors mediate interference effects in the RI-style tasks (see Figure 2 for examples of both accounts). Furthermore, they challenge the view that WM consolidation entails access to an immutable central processing bottleneck that processes a wide variety of information and tasks, one at a time (Tombu et al., 2011). Because these results contradict foundational research, understanding the mechanisms behind them will have a significant impact on accounts of WM consolidation going forward.

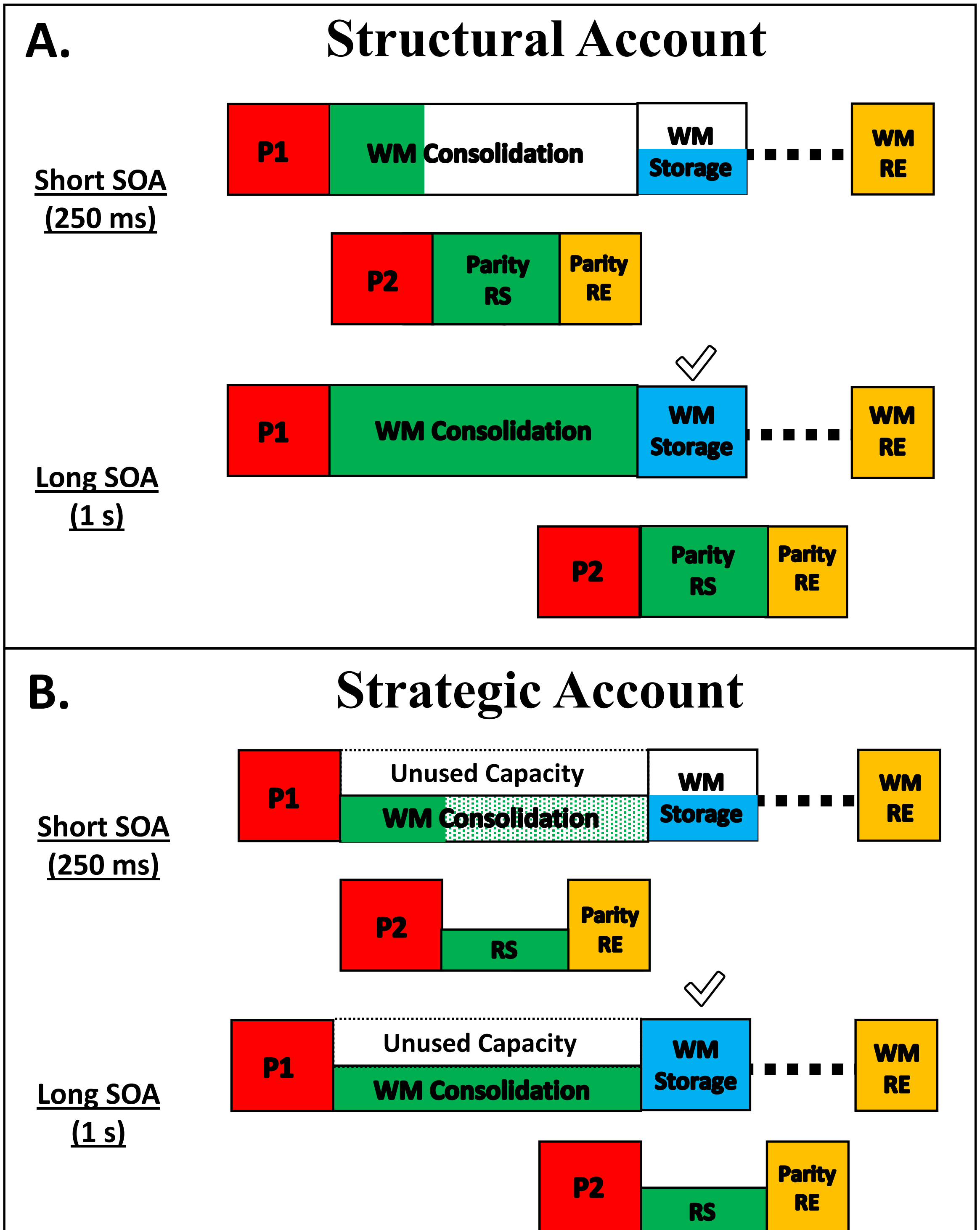

Figure 2. A depiction of the structural (A) vs. strategic (B) accounts. The red boxes represent visual perception of either stimulus 1 (P1), or stimulus 2 (P2). The green box represents either WM consolidation of P1, or T2 parity response selection (RS) for P2. The length of the green rectangle indicates the duration of WM consolidation, or duration of RS of P2, whereas the height represents all available capacity. The blue box represents the amount of information that was successfully stored into WM. The yellow box represents either the response execution for the WM task (WM RE), or response execution for the T2 parity judgment (Parity RE).

In the present study, Nieuwenstein and Wyble's RI results were replicated and extended, including demonstrating robustness to varying WM response demands. The RI effect was obtained not only for verbal stimuli (letters), but also for non-verbal visual stimuli (colors) that are widely used in visual WM storage capacity experiments, easing the future integration of results from storage and consolidation rate studies. It was then determined that slow estimates of WM consolidation are not a result of strategic prioritization of the 2AFC task over WM encoding, supporting the structural account. However, it was found that the RI effect was abolished when a response to the second task was deferred until the end of the trial suggesting a critical role of the speeded nature of the 2AFC. Moreover, a sensorimotor decision and response—not mere distraction—were required to obtain RI.

EXPERIMENTS 1 AND 2

Information that has been successfully stored into WM can be used and manipulated to serve the needs of an ongoing task for a short period of time, typically lasting only for a matter of seconds (Baddeley, 2012; Baddeley & Hitch, 1974; Phillips, 1974). The task of Nieuwenstein and Wyble (2014) required participants to report all four letters on each trial by using the corresponding keys on a computer keyboard, potentially leading to diminished WM performance over the course of each trial's report period due to accumulating time and interference from reporting prior letters. This motivated to determine if only such whole-report WM tasks are vulnerable to RI, or if this effect also extends to less demanding report tasks, such as change detection. A change detection response requires participants to determine if a change has been made from the initial WM sample to the subsequent WM probe. Instead of having participants input each letter (e.g., type in each of the 4 letters shown), a change detection task requires participants to use one of two keys to identify whether a new stimulus (i.e., showing a second set of four letters) is the "same" or "different" from the original array. Change detection can be seen as a less demanding report as a single response will be made using the same two keys, versus whole-report, where if letters were selected from the entirety of the alphabet, there can potentially be 26 different key options. Moreover, change detection responses do not demand serial response-selection and motor acts associated with typing four letters; each such event entails the possibility of interference with memory maintenance. In addition, using change detection will shorten the response times compared to whole-report. This is important because it will decrease the chance for decay of the later-reported items in the memory array. Potential prolonged response times to a WM probe could lead to time-based WM decay of already-consolidated items masquerading as consolidation of fewer items into WM.

In two experiments, it was examined whether the longer estimate of WM consolidation found using the retroactive dual-task interference paradigm (Nieuwenstein & Wyble, 2014) compared to other measurement approaches is a result of the nature of WM report. First, Experiment 1 was designed to replicate the original RI effect found by Nieuwenstein and Wyble (2014). In Experiment 2, instead of having participants report each letter, a less demanding change detection response was required. This task manipulation allowed the investigation of robustness of RI to varying WM response demands. All other aspects of the original Nieuwenstein and Wyble (2014) paradigm, such as timing and stimuli, remained the same.

Figure 3 includes the design of both Experiments 1 and 2. For Experiment 1, participants were asked to encode a string of four letters visually displayed on the computer screen, and later, after a variable delay, recalled the original four letters by using the corresponding keys. During the delay period, participants either did or did not also see a visual mask (to disrupt iconic memory, and potentially, WM consolidation) and/or a digit (for 2AFC parity judgment). The only difference from Experiment 1 to Experiment 2 is that instead of recalling the to-be-remembered letters, participants were presented with another set of letters, and asked whether the letter string was the same or different (change detection).

Experiment 1 Method

Participants. Data from 16 undergraduate students (5 males; 18-32 years old, mean (M) age = 22 years, SD = 3.51) were collected. Participants were at least 18 years of age, did not report any vision problems, and reported no history of neurological problems. Participants were compensated course credit for their completion of the task.

Materials. The experiment was designed and run in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007). Stimuli were displayed on a CRT monitor set to 1,600 X 1,200 pixels, with a 70 Hz refresh rate, driven by a Linux-based Dell computer.

Similar to Nieuwenstein and Wyble (2014), Experiment 1 required recall of a string of four letters randomly selected without replacement from the English alphabet. The letters M, W, and all vowels were excluded from the pool of letters. This was done to exclude the chance of potential real words from being shown. Pound symbols ("#") were used to mask the letters on masked trials. All stimuli were presented in the center of the screen on a gray background. A 20-point Arial font was used for letters and digits. A 24-point boldfaced Arial font was used for the masks.

Design and procedure. A within-subject factorial design was used for Experiment 1, with SOA, the presence of a mask, and the presence of a second task as factors. The SOA, or the time from the start of the WM-sample (4 letters) to the start of the parity judgment task (T2), was either 250, 500, or 1,000 ms. The trials were randomly intermixed, leading to 240 total trials (12 conditions; 20 trials per condition).

The design of Experiment 1 is displayed in Figure 3a. Each trial began with four placeholder crosses in the center of the screen. Participants initiated the trial by pressing the spacebar. The four crosses remained in the center of the display for 400 ms. The WM sample then took the place of the crosses and was shown for 100 ms. Half the trials presented four "#" symbols for 100 ms to mask the WM sample. For the other half of the trials that did not include the mask, there was a blank screen for a 100 ms interval. This was then followed by a 50, 300, or 800 ms blank screen interval until the presence/absence of a 2AFC parity judgment task. For the dual-task trials, a digit was then shown for 100 ms. Participants were to respond as fast as possible using the ",<" or ".>" key to identify whether the digit shown was odd or even, respectively. After their 2AFC digit parity response, they immediately were presented with the WM probe. Specifically, they were prompted to input, using the corresponding keys, which four letters were originally shown at the beginning of the trial. For the single-task trials, there was a 100 ms blank interval when the digit would have been shown. There was then a 1000 ms retention interval before the participants were prompted to input the four letters that were originally shown.

General analysis pipeline. Repeated-measures analysis of variance (ANOVA) was used to evaluate WM performance (T1). RI with WM consolidation is detected by a significant interaction of SOA by the presence of a second task. A reasonable alternative approach to examining whether RI with WM consolidation is obtained is by a main effect of SOA on dual-task trails alone (i.e., excluding single-task trials from analysis). However, single-task trials were included to look for an interaction, both to maintain comparability to Nieuwenstein and Wyble (2014), and to provide an inbuilt control for time-based decay of WM. A Bayesian Repeated Measures ANOVA was also conducted using the JASP statistics package (JASP Team, 2018). The Bayes factors (BF) were used to quantify the ratio of evidence in favor or against the inclusion of each factor in the model (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017; Vandekerckhove, Rouder, & Kruschke, 2018; Wagenmakers et al., 2018). A BF < 1 represents evidence against inclusion of a factor or interaction, whereas a BF > 1 represents evidence in favor of inclusion of a factor or interaction. A BF of 0.1 is interpreted as a ratio of 10:1 against the inclusion, whereas a BF of 10 is interpreted as a 10:1 ratio for the inclusion of the factor (see Doorn et al., 2019).

Experiment 1 Results

WM Task (T1). A repeated-measures analysis of variance (ANOVA) was used to examine the total number of correctly recalled letters including SOA, the presence of a mask, and the presence of a second task (parity judgment) as factors. See Figure 3C for a graph of the WM results. There were significant main effects of all factors (all *p*s $< .001$, η_p^2s $> .560$), and a significant interaction of SOA and the presence of a second task, *F*(2,

30) = 23.57, $p < .001$, $\eta_p^2 = 0.61$. The current findings successfully replicated RI of a parity task (T2) on WM consolidation of Nieuwenstein and Wyble (2014).

Parity Judgment Task (T2). The overall mean accuracy on the parity judgment task was .90 (SD = .11). There was no significant main effect of SOA on parity judgment response accuracy, $F(2, 30) = 3.61$, $p = .07$, $\eta_p^2 = 0.19$. The analysis of T2 accuracy also revealed no significant main effect of mask presence, nor an interaction of SOA and the presence of a mask (both *p*s > .43). Table 1 includes T2 mean accuracy and standard deviations across experiments and conditions.

Response times for T2 were also analyzed (see table 2 for full T2 mean RTs across experiments). There was a significant effect of SOA on RTs, $F(2,30) = 5.50$, $p < .01$, $\eta_p^2 = 0.27$. Response times at SOAs 250 and 500 were slower than at SOA 1,000, consistent with a PRP evoked by WM encoding slowing down a subsequent sensorimotor decision (Jolicoeur & Dell'Acqua, 1998). There was not a significant main effect of mask presence nor an interaction of SOA and mask presence (both *p*s > .18).

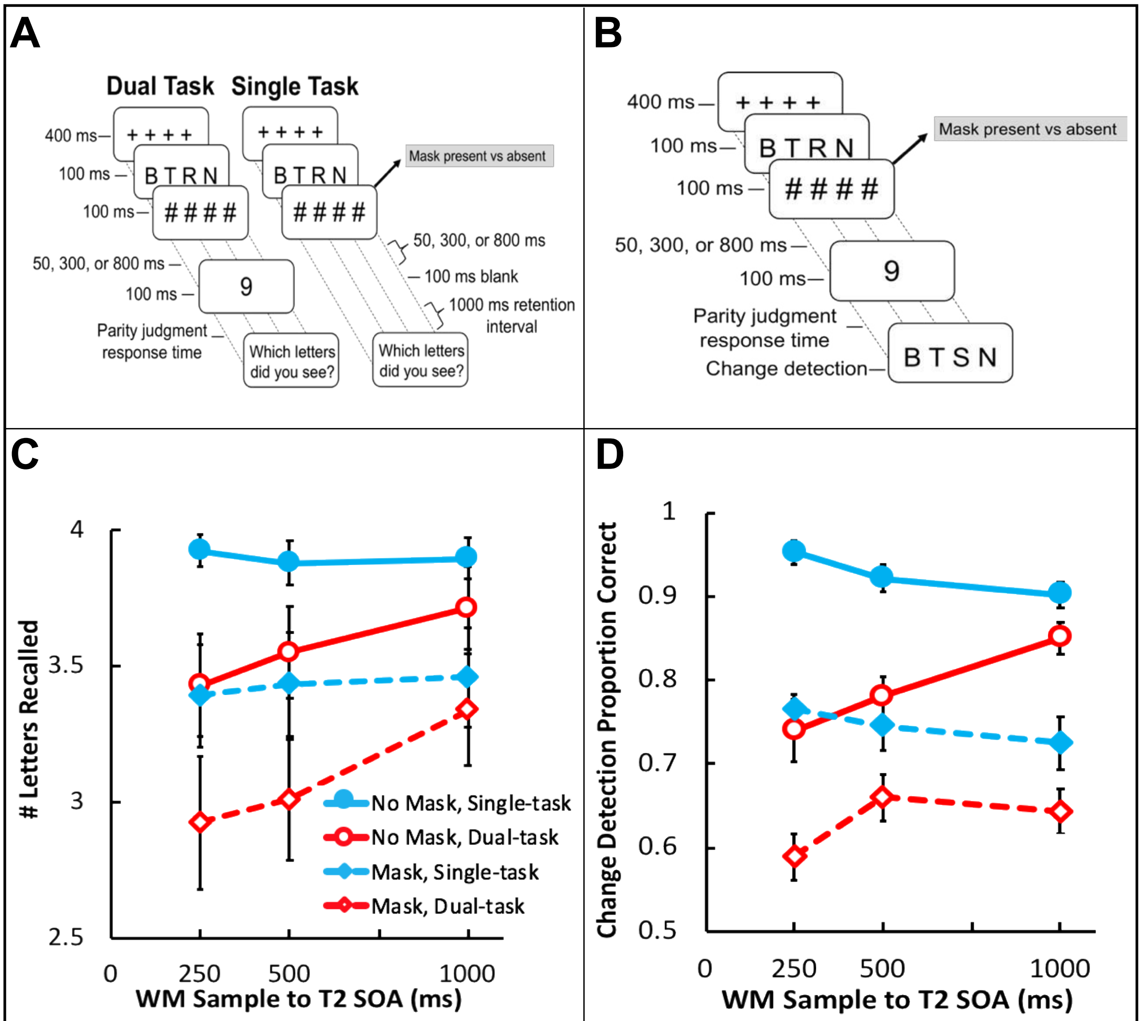

Figure 3. Design and results for Experiments 1 and 2. Error bars show the standard error of the mean.
A) Single- vs dual-task trials for Experiment 1.
B) Dual-task trials for Experiment 2. Even though not depicted in this figure, Experiment 2 does include single-task trials adapted to the change detection manipulation.
C) Experiment 1 WM performance.
D) Experiment 2 WM performance.

Table 1
T2 mean accuracy (& standard deviation) across experiments and conditions

	150 ms	250 ms	500 ms	1000 ms
Experiment 1 (Whole-report)				
Unmasked		0.916 (0.068)	0.919 (0.083)	0.828 (0.231)
Masked		0.931 (0.079)	0.906 (0.134)	0.791 (0.272)
Experiment 2 (Change detection)				
Unmasked		0.831 (0.212)	0.838 (0.172)	0.866 (0.159)
Masked		0.897 (0.138)	0.875 (0.106)	0.856 (0.153)
Experiment 3 (Color WM sample)				
Unmasked		0.923 (0.089)	0.899 (0.163)	0.950 (0.067)
Masked		0.938 (0.059)	0.928 (0.082)	0.930 (0.060)
Experiment 4 (Deferred T2)				
Unmasked		0.916 (0.068)	0.919 (0.077)	0.897 (0.096)
Masked		0.853 (0.134)	0.878 (0.098)	0.922 (0.066)
Experiment 5 (Deferred T2 + instructed priority)				
Unmasked (WM Prioritized)		0.888 (0.090)	0.884 (0.123)	0.941 (0.058)
Masked (WM Prioritized)		0.831 (0.115)	0.822 (0.126)	0.944 (0.051)
Unmasked (T2 Prioritized)		0.909 (0.117)	0.922 (0.118)	0.925 (0.105)
Masked (T2 Prioritized)		0.925 (0.111)	0.919 (0.101)	0.934 (0.111)
Experiment 6 (Instructed priority)				
Unmasked (WM Prioritized)		0.972 (0.048)	0.969 (0.068)	0.969 (0.054)
Masked (WM Prioritized)		0.931 (0.073)	0.953 (0.059)	0.966 (0.054)
Unmasked (T2 Prioritized)		0.984 (0.030)	0.975 (0.032)	0.991 (0.020)
Masked (T2 Prioritized)		0.978 (0.052)	0.978 (0.055)	0.969 (0.036)
Experiment 7 (T2: 2nd WM task)				
Unmasked	0.922 (0.074)	0.880 (0.126)	0.898 (0.079)	0.965 (0.064)
Masked 1	0.938 (0.060)	0.930 (0.094)	0.895 (0.101)	0.949 (0.073)
Masked 2	0.922 (0.084)	0.840 (0.116)	0.809 (0.156)	0.934 (0.066)
Masked (1 & 2)	0.836 (0.118)	0.867 (0.140)	0.852 (0.075)	0.961 (0.050)
Experiment 8 (Change detection)				
Unmasked	0.884 (0.087)	0.900 (0.084)		0.853 (0.131)
Masked	0.878 (0.134)	0.863 (0.126)		0.881 (0.138)
Experiment 9 (Stimulus detection)				
Unmasked		0.974 (0.038)	0.959 (0.058)	0.972 (0.053)
Masked		0.964 (0.051)	0.947 (0.055)	0.942 (0.084)

Note: The mean values are reported for each condition, with standard deviation reported in parenthesis. Values are omitted for some experiments due to not collecting data for that SOA.

Table 2

T2 RTs across experiments and conditions. RTs reported in seconds.

	150 ms	250 ms	500 ms	1000 ms
Experiment 1 (Whole-report)				
Unmasked		1.546 (0.653)	1.430 (0.672)	1.254 (0.522)
Masked		1.488 (0.630)	1.270 (0.409)	1.262 (0.410)
Experiment 2 (Change detection)				
Unmasked		1.049 (0.387)	1.001 (0.467)	0.948 (0.305)
Masked		1.000 (0.415)	0.974 (0.352)	0.967 (0.341)
Experiment 3 (Color WM sample)				
Unmasked		1.137 (0.356)	1.892 (1.916)	1.061 (0.325)
Masked		1.103 (0.300)	1.370 (0.877)	1.077 (0.308)
Experiment 6 (Instructed priority)				
Unmasked (WM Prioritized)		1.599 (0.487)	1.449 (0.479)	1.428 (0.476)
Masked (WM Prioritized)		1.512 (0.349)	1.348 (0.332)	1.331 (0.357)
Unmasked (T2 Prioritized)		1.496 (0.527)	1.424 (0.417)	1.277 (0.398)
Masked (T2 Prioritized)		1.501 (0.659)	1.378 (0.489)	1.313 (0.425)
Experiment 8 (Change detection)				
Unmasked	1.021 (0.293)	1.089 (0.316)		1.063 (0.367)
Masked	1.194 (0.546)	1.252 (0.618)		1.178 (0.427)
Experiment 9 (Stimulus detection)				
Unmasked		0.721 (0.250)	0.705 (0.272)	0.611 (0.228)
Masked		0.757 (0.362)	0.710 (0.267)	0.763 (0.431)

Note: Only experiments with an immediate speeded T2 response were included in this table, with standard deviation reported in parenthesis.

Experiment 2 Method

Participants. Sixteen undergraduate students (4 males; 18-30 years old, M = 22.1 years, SD = 4.08) completed Experiment 2. They were compensated course credit for their time.

Materials. The same setup used in Experiment 1 was used in Experiment 2. In Experiment 1 participants had to recall the letters for T1. However, in Experiment 2, participants were to provide a recognition, change detection response. Participants were shown four letters at the end of the trial and had to determine if the new letters were the same as the original WM sample. The four letters shown at the end of the trial were selected from the same pool of letters used for the WM sample. The proportion of no-change vs change trials was set at 50%. For no-change trials, all four letters were the same as in the sample array, with the exact same order. For change trials, only one letter was different, and the other three letters were the same, in the exact position. The position of the changed letter was distributed uniformly among the letter positions.

Design and procedure. The experimental design and procedure of Experiment 2 were identical to that of Experiment 1 except that instead of being prompted "Which letters did you see?" and inputting the four letters, participants were just shown a new array of four letters that remained on the screen until a response was made. They used the "Z" (no change) or "X" (change) key to respond to the change detection portion of the task. Figure 3B shows an example of the dual-task portion of the task.

Experiment 2 Results

Memory Task (T1). Similar results to Experiment 1 were found in Experiment 2. Change detection performance was examined using a repeated-measures ANOVA including SOA, the presence of a mask, and the presence of a second task as factors. There was a non-significant main effect of SOA, $F(2, 30) = 1.06$, $p > .36$, $\eta_p^2 = 0.07$. There were significant main effects revealed for the presence of a mask, and presence of a second task (both *p*s < .001, η_p^2s > .88). There was also a significant interaction of SOA and the presence of a second task, $F(2, 30) = 11.08$, $p < .001$, $\eta_p^2 = 0.43$. Figure 3D includes a graph of the WM results for Experiment 2.

Parity Judgment Task (T2). There was an average T2 performance of 86.0% (SD = .16). See Table 1 for a full list of mean accuracies from T2 performance across conditions. The analysis of T2 response accuracy revealed no significant effects of SOA, or mask presence (both *p*s > .33). There was not a significant interaction between SOA and the presence of a mask, $F(2, 30) = 1.95$, $p = .16$, $\eta_p^2 = 0.12$.

T2 response times were also analyzed and revealed no significant effects of SOA, presence of a mask, or interaction of the two (all *p*s > .22). See Table 2 for T2 mean RTs across conditions.

Between-subjects analysis (Experiment 1 & 2). The WM task data from Experiments 1 and 2 were analyzed together. By doing so, the N-items results from Experiment 1 (whole-report) were converted to the percent correct needed to match Experiment 2. This was done by dividing the items recalled for each condition by 4 (the number of letters present). For instance, if a participant on average recalled 3 items for a given condition, this would equate to 75%. After the conversion from N-items recalled to percent correct, WM accuracy was analyzed using a mixed ANOVA including SOA, presence of a second task, and presence of a mask as within-subject factors, and response demand (recall vs. change detection) as a between subject factor. The results did not yield a significant SOA x Presence of a second task x Response demand (recall vs change detection) interaction, $p = .41$, BF = .135 (~7:1 evidence against the inclusion of the interaction). See Table 3 for full results.

Table 3
Experiments 1 & 2 (WM accuracy) repeated measures ANOVA with response demand as a between-subject factor

	F	df	P	η_p^2	BF
SOA	9.723	2,60	<.001	.245	6.196
Presence of a second task	115.797	1,30	<.001	.794	$7.40e^{29}$
Mask	180.159	1,30	<.001	.857	$1.11e^{50}$
Response demand	8.766	1,30	.006	.226	7.671
SOA x Response demand	2.848	2,60	.066	.087	2.86
Presence of a second task x Response demand	1.216	1,30	.279	.039	.421
Mask x Response demand	3.188	1,30	.084	.096	2.525
SOA x Presence of a second task	26.853	2,60	<.001	.472	92620.5
SOA x Mask	.110	2,60	.896	.004	.059
Presence of a second task x Mask	.156	1,30	.695	.005	.175
SOA x Presence of a second task x Response demand	.906	2,60	.410	.029	.135
SOA x Mask x Response demand	2.491	2,60	.091	.077	.295
Presence of a second task x Mask x Response demand	.412	1,30	.526	.014	.307
SOA x Presence of a second task x Mask	.487	2,60	.617	.016	.151
SOA x Presence of a second task x Mask x Response demand	.840	2,60	.437	.027	.146

Experiments 1 and 2 Discussion

Experiment 1 was a successful replication of RI on WM consolidation (Nieuwenstein & Wyble, 2014). Participants were shown four letters and asked to remember them for later recall. After the presentation of the letters, on half of trials, a digit appeared and participants made a speeded response to identify whether it was odd or even. The number of correctly-reported WM items was reduced by performance of a second (digit parity 2AFC) task, but this RI effect attenuated with WM sample-parity task SOA. Experiment 2 demonstrated that WM response demands do not explain the results. Specifically, when shifted from a whole-report to a change-detection paradigm, a significant interaction between delay duration and the presence of a second task, was still observed exactly as in whole-report (Experiment 1). Moreover, this interaction did not differ between Experiments 1 and 2, suggesting that RI with WM consolidation is invariant to WM report demands. The reported T2 RT data will be analyzed across experiments and discussed in the later sections.

EXPERIMENT 3

All of the previous experiments used letters for the WM sample. Letters are a form of verbal visual stimuli. This means the WM sample items can be rehearsed (aloud or subvocally) and remembered. However, does the RI effect translate to non-verbal visual stimuli? Experiment 2 from Nieuwenstein and Wyble (2014) showed RI with WM consolidation is robust when using complex spatial stimuli (Kanji characters) as the memory array. However, past research has shown WM is sensitive to the visual complexity of stimuli (Alvarez & Cavanagh, 2004; Alvarez & Cavanagh, 2008; Eng, Chen, & Jiang, 2005; Hao, Becker, Ye, Liu, & Liu, 2018; Mance, Becker, & Liu, 2012; Miller, Becker, & Liu, 2014; Song & Jiang, 2006). For instance, it was found that more objects can be remembered for simple visual features (e.g., color) versus complex visual form, such as Chinese characters (Alvarez & Cavanagh, 2004). Therefore, the differences in WM capacity for the various stimulus types may lead to fluctuating results. Thus, the goal of Experiment 3 was to replicate Experiment 2 while having participants encode colored boxes as the WM sample. This arrangement was to test whether the RI with WM consolidation effect translates to a more strictly visual form of visual WM. Also, since much or most visual WM research uses simple single-feature items such as color patches or oriented bars, Experiment 3 allowed for the current results to be better related to the existing body of literature.

Before completion of the discussed task, participants performed a k-estimation task lasting approximately 10 minutes. This task provided an estimate of the participant's WM storage capacity by using a standard whole-array probe color change detection task (c.f., Luck & Vogel, 1997). These data were collected for a future goal of looking at individual differences to elucidate how WM capacity relates to WM consolidation. However, as this goal is not the focus of the current thesis, there will be no further mention of this task.

Method

Participants. Data from 16 undergraduate students (5 males; 18-37 years old, M = 23.3 years, SD = 4.81) were included in the analysis. Participants were compensated course credit for their time.

Materials, design, and procedure. Experiment 3 used the same setup as Experiment 2. However, in Experiment 3 participants had to remember the colors of four boxes rather than the identity of four letters. Participants were then shown a new array of four colored boxes at the end of the trial, and had to determine if they were all the same as the original WM sample. The four colored boxes shown at the end of the trial were selected from the same pool of colors selected from the WM sample. The pool consisted of 9 different colors. "No-change" trials used all the same colored boxes from the original memory array. However, "change" trials selected at random from one of the remaining 5 colors not used in the WM sample, and replaced the color of one of the four boxes. The 9 possible colors presented were black, red, green, blue, yellow, cyan, purple, pink, and orange. Instead of using a pound symbol as a mask, which is adequate for letters, but not colors, each element of the mask in this experiment (see Figure 10) consisted of a 3-by-3 grid, with each sub-element randomly filled with one of the 9 possible colors, without replacement. The whole mask consisted of four such elements placed in the same locations as the four color patches of the WM sample. Identical to Experiment 2, the proportion of no-change vs change was set evenly at 50%. For no-change trials, all four boxes were the same, and in the same order. For change trials, only one box was different, and the other three colors were the same, in the same order. The position of the different colored box for the change trials were evenly distributed among the four letter positions. Timing was exactly the same as Experiment 2. The task design for Experiment 3 can be found in Figure 4.

Analyses. In addition to an ANOVA incorporating all factors in the design of Experiment 3, it was planned *a priori* to separately consider masked and unmasked trials, because there was uncertainty about whether the 100 ms sample-mask SOA would result in a floor effect (chance performance) for the masked trials. Support for this view comes from Vogel et al. (2006), who estimated that only approximately two items should be successfully encoded with these timing parameters. Nevertheless, the chosen timing was used in order to more directly replicate Nieuwenstein and Wyble (2014). The concern that the mask would lead to floor performance justifies the separate analysis of these conditions because including conditions with floor performance could suppress SOA x dual-task effects in the ANOVA that considers the entire task design.

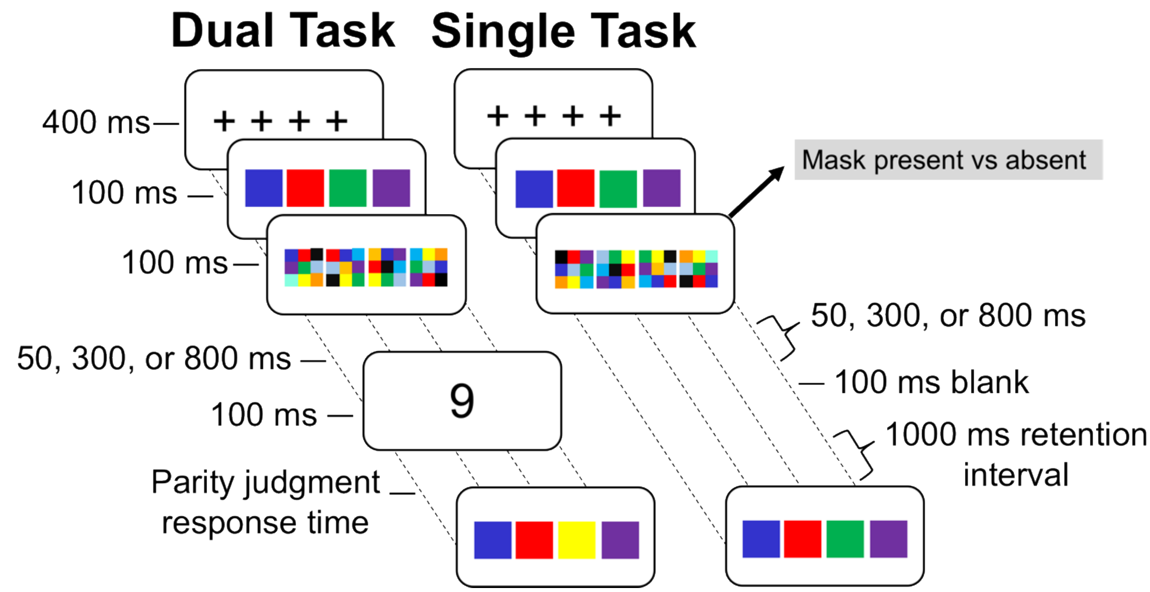

Figure 4. Single- and dual-task trials for Experiment 3. The dual-task trial shown is an example trial for a "change" response. The single-task trial shown is an example of a "no-change" response.

Results

Memory Task (T1). A repeated-measures analysis of variance (ANOVA) was used to examine the change detection performance including SOA, the presence of a mask, and the presence of a second task (parity judgment) as factors. There was a non-significant main effect of SOA, $F(2,30) = 1.85$, $p = 0.17$, $\eta_p^2 = 0.11$. There were significant main effects revealed for the presence of a mask, and presence of a second task ($p < .001$ for both effects). There was also a significant interaction of SOA and the presence of a second task, $F(2, 30) = 5.17$, $p = .01$, $\eta_p^2 = 0.26$.

However, it must be noted that there was an a priori reason to analyze masked and unmasked trials separately (see *Planned analyses*, above). When running a repeated-measures ANOVA with SOA and presence of a second task as factors with the unmasked data a significant SOA x presence of a second task interaction was obtained, $F(2,30) = 3.514$, $p = 0.04$, $\eta_p^2 = 0.19$. There was also a significant main effect of the presence of a second task, $F(2,30) = 17.44$, $p < 0.001$, $\eta_p^2 = 0.54$. There was a non-significant main effect of SOA ($p > 0.19$).

For the masked data, there was a significant interaction of SOA x the presence of a second task, $F(2,30) = 4.484$, $p = 0.02$, $\eta_p^2 = 0.23$. There was not a significant main effect of SOA, nor the presence of a second task (both *p*s > 0.12). Results for the WM task are shown in Figure 5.

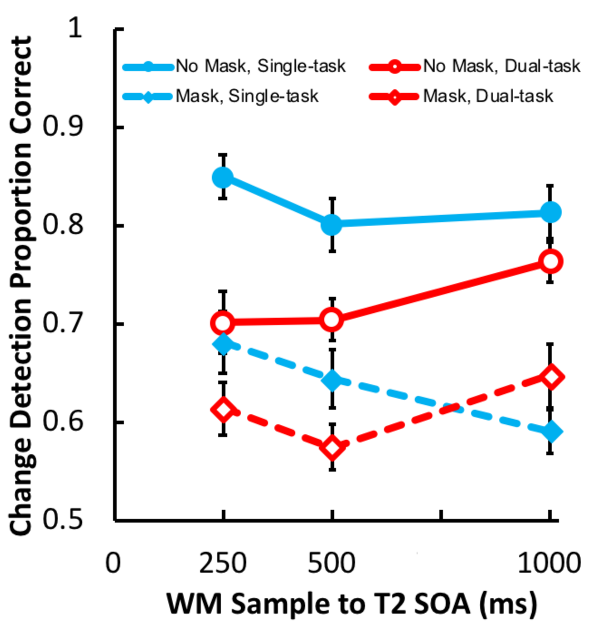

Figure 5. Results for WM performance. Error bars show the standard error of the mean.

Parity Judgment Task (T2). There was an average T2 performance of 93% (SD = .09; see Table 1 for all results). The analysis of T2 response accuracy revealed no significant main effects or interaction of SOA and mask (all *p*s > .37 and η_p^2s < 0.07).

T2 response times were also analyzed and revealed a significant main effect of SOA, $F(2, 30) = 4.45$, $p = .05$, $\eta_p^2 = 0.23$. There was no significant main effect of the presence of a mask, or interaction of SOA by presence of a mask revealed (both ps > .88 and η_p^2s < 0.002). See Table 2 for a list of all T2 RT averages.

Between-subjects analysis (letters vs colors). Next, Experiment 2 and Experiment 3 were analyzed together including stimulus type (letters vs. colors) as a between subject factor. The ANOVA revealed a non-significant SOA x presence of a second task x stimulus type interaction (p = .44, BF = .136, ~7:1 evidence against the inclusion of the interaction), consistent with similar consolidation rates/RI effects for letter and color-patch stimuli. See Table 4 for full results.

Table 4
Experiments 2 & 3 (WM accuracy) repeated measures ANOVA with stimulus type as a between-subject factor

	F	*df*	*P*	η_p^2	*BF*
SOA	.758	2,60	.473	.025	.049
Presence of a second task	87.801	1,30	<.001	.745	$4.84e^{15}$
Mask	214.504	1,30	<.001	.877	$3.77e^{42}$
Stimulus type	17.16	1,30	.001	.364	56.366
SOA x Stimulus type	2.306	2,60	.108	.071	.376
Presence of a second task x Stimulus type	9.733	1,30	.004	.245	24.091
Mask x Stimulus type	1.136	1,30	.295	.036	.405
SOA x Presence of a second task	13.798	2,60	<.001	.315	5157.297
SOA x Mask	1.325	2,60	.274	.042	.159
Presence of a second task x Mask	4.153	1,30	.050	.122	2.453
SOA x Presence of a second task x Stimulus type	.834	2,60	.439	.027	.136
SOA x Mask x Stimulus type	.283	2,60	.755	.009	.119
Presence of a second task x Stimulus type	1.261	1,30	.270	.040	.325
SOA x Presence of a second task x Mask	.140	2,60	.870	.005	.104
SOA x Presence of a second task x Mask x Stimulus type	1.933	2,60	.154	.061	.721

Discussion

Experiment 3 aimed to determine whether RI with WM consolidation would be obtained using non-verbal visual stimuli rather than verbal visual stimuli. The collected data suggest that the RI effect obtain using verbal visual stimuli such as letters, also translates to non-verbal visual stimuli such as colored-boxes. Moreover, the RI effect did not differ between stimulus types (letter vs. colors).

One may argue that colored-boxes may be verbally coded, rather than visually coded. Past studies have used articulatory suppression by having participants complete a WM task (colored-boxes sample) with the presence of a concurrent verbal load and found no significant difference in VWM capacity than with no verbal load present (Vogel, Woodman, & Luck, 2001; Sense, Morey, Prince, Heathcote, & Morey, 2016). These findings suggest there is little influence of verbal WM when using colored-boxes as the WM sample supporting the idea that Experiment 3 is tapping into VWM mechanisms.

EXPERIMENT 4

Experiments 1 and 2 confirmed that the RI on WM consolidation reported by Nieuwenstein and Wyble (2014) was not an artifact of using a whole-report WM probe. Thus, Experiment 4, investigated why a slow consolidation process (lasting approximately 1s) was observed by Nieuwenstein and Wyble (2014), while other research has suggested a much faster consolidation process. Experiment 4 addressed whether slow consolidation was an ever-present phenomenon—a structural constraint of the WM system—to which prior measurement techniques were insensitive, or if slow consolidation was instead a result of a strategic slowing of WM processing in order to better accomplish a second task (digit parity 2AFC)—which would suggest a prolongation of an otherwise rapid consolidation process restricted to the dual-task trials of the Niewenstein and Wyble (2014) task. Thus, the goal of Experiment 4 was to determine whether slow consolidation might be driven by the relative prioritization of the speeded 2AFC judgment (T2) over the unspeeded WM task. One explanation is that the immediate, speeded response required by

T2 could implicitly drive participants to assign it high priority. Given limited processing resources (Kahneman, 1973; Koch et al., 2018; Marois & Ivanoff, 2005; Tombu & Jolicoeur, 2003), increasing the priority of T2 could result in decreased resources—leading to slower or queued performance—for WM consolidation (T1) (cf., Schumacher et al., 2001). Thus, Experiment 4 used a task manipulation to determine whether an unspeeded T2 (parity judgment) with a deferred response would still lead to a robust RI effect. For the remainder of the study a change detection design was used instead of whole-report for WM, which as noted, reduces a drawn-out report process and reduces it to a punctate event.

For establishing the unspeeded condition, the parity judgment response was deferred until after participants responded to the WM change detection task. Having participants respond to T2 at the end of the trial should minimize the potential implicit prioritization of the parity task over the WM task. The hypothesis is that the slow estimate of WM consolidation (assessed by RI) was caused by the low relative priority of unspeeded, delayed WM reports compared to speeded, immediate parity reports—in other words, that the presence of a high-priority T2 led to a strategic prolongation of WM consolidation (see Figure 2b).

Method

Participants. Data from 16 undergraduate students (16 females; 18-41 years old, M = 21.4 years, SD = 5.54) were collected. Participants were compensated course credit for their completion of the task.

Materials. The same setup used in Experiment 2 was used in Experiment 4.

Design and procedure. Experiment 4 followed a similar design as Experiment 2. The difference between Experiments 2 and 4 was that for the dual-task trials, the parity judgment task was unspeeded (Figure 6). Instead of having the participants respond as fast as possible, they were to hold their response until the end of the trial, after they responded to T1. For dual-task trials, since the parity task was unspeeded, the additional delay after the appearance of the digit and before the WM probe was set to 1000 ms, which was done to match the setup of the single-task trials.

Results

Memory Task (T1). A repeated measures ANOVA of WM performance revealed no main effect of SOA ($F(2, 30) = 1.70$, $p = .21$, $\eta_p^2 = 0.10$), nor an interaction of SOA and presence of a second task, $F(2, 30) = 2.57$, $p = .11$, $\eta_p^2 = 0.15$. There were significant main effects of both the presence of a second task and the presence of a mask (both ps < 0.003, η_p^2s > 0.46).

Parity Judgment Task (T2). There was an average T2 performance of 90% (SD = .09; see Table 1 for mean T2 accuracy across conditions). The analysis of T2 response accuracy revealed no significant effects of SOA, nor mask presence (both ps > .09). There was a significant interaction between SOA and the presence of a mask, $F(2, 30) = 3.873$, $p = .03$, $\eta_p^2 = 0.21$. T2 response times were not analyzed because the task required a deferred response.

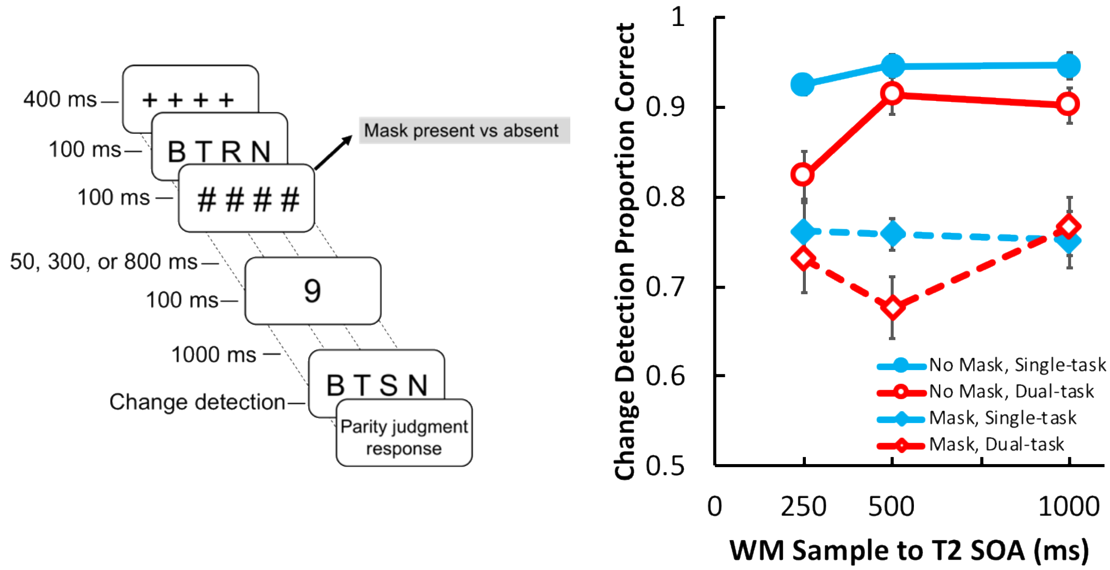

Figure 6. (Left) Dual-task trials for Experiment 4. (Right) Working memory performance. Error bars show the standard error of the mean.

Discussion

The hypothesis was that obtaining RI instead of PI might be driven by the relative prioritization of the speeded parity judgment over the unspeeded WM task. Thus, Experiment 4 decreased parity judgment priority by making the responses unspeeded. This change in task prioritization abolished the interaction between delay duration and the presence of a second task, consistent with the original hypothesis. Although, as hypothesized, the interaction of SOA by presence of T2 was not statistically significant, it must be acknowledge that the results were marginal (p = 0.11, BF = 0.42) and did not strongly support the original hypothesis that high T2 priority drove prolonged WM consolidation. Thus, Experiments 5 and 6 address relative prioritization by using direct manipulations of task priority to more strongly test the prioritization account.

EXPERIMENT 5

Marginal results found in Experiment 4, which used a weak manipulation of task priority, motivated the idea to provide a stronger manipulation of priority in Experiment 5. Past research has used monetary (Capa, Bouquet, Dreher, & Dufour, 2013; Rieger, Mittelstädt, Dignath, & Kiesel, 2019; Schevernels, Krebs, Santens, Woldorff, & Boehler, 2014; Schumacher et al., 2001; Zedelius, Veling, Bijleveld, & Aarts, 2012) and non-monetary (Erkal, Gangadharan, & Koh, 2018; Greenhouse & Wessel, 2013; Ye et al., 2017) reward incentives to successfully manipulate task preparation. Thus, to cause a stronger manipulation of task priority, Experiment 5 adapted the task used in Experiment 4, but explicitly manipulated the priority of the two tasks between participants by providing feedback on performance. This change provided explicit prioritization to reordered tasks. Specifically, one task was designated as high priority (WM or parity); for this task participants earned or lost 500 points for correct or error responses, respectively. In addition, high priority task errors resulted in a 5s delay before the next trial (influencing motivation). For the low priority task (task not assigned to high priority) however, if participants correctly responded to the task they received 10 points. If participants incorrectly responded they lost 10 points. Performance on the low priority task did not affect timing between trials.

In Experiment 5a, the WM task was given high priority, and the parity judgment task was given low priority. In Experiment 5b, the priority of each task was flipped. Besides the low and high priority manipulations, the task remained identical to Experiment 4. The hypothesis was that this stronger manipulation of task priority would abolish RI with WM consolidation effect.

Method

Participants. Data from 16 undergraduate students (Experiment 5a: 4 males, 18-26 years old, M = 20.6 years, SD = 1.82; Experiment 5b: 5 males, 18-25 years old, M = 21.1, SD = 2.78) were collected for each experiment. Participants were compensated course credit for their completion of the task.

Materials, design, and procedure. The design of Experiments 5a and 5b were similar to that of Experiment 4; the only difference was to provide higher priority by giving/taking points dependent upon performance, as well as affecting timing of the inter-trial interval. At the conclusion of each trial, when a participant correctly or incorrectly responded to the higher priority task, the screen turned green or red, respectively. If there was a green screen, participants earned 500 points, and were allowed to move on to the next trial immediately. However, if they received a red screen, they lost 500 points, and had a 5 second delay until they were able to advance to the next trial. Participants were given a running total of their points after each trial and a grand total at the end of the task.

Experiment 5a (WM Prioritized) Results

Memory Task (T1). A repeated measures ANOVA of WM performance revealed no main effects of SOA and the presence of a second task (both ps > 0.17, η_p^2s < 0.12), nor an interaction of SOA and presence of a second task, $F(2, 30) = 1.83$, $p = .18$, $\eta_p^2 = 0.11$. There was a significant main effect of the presence of a mask, $F(1, 30) = 100.88$, $p < .001$, $\eta_p^2 = 0.87$.

Parity Judgment Task (T2). There was an average T2 performance of 88.5% (SD = .11; see Table 1 for full list of T2 mean accuracies). The analysis of T2 response accuracy revealed a significant main effect of SOA, $F(2, 30) = 7.37$, $p = .008$, $\eta_p^2 = .33$. A significant main effect of masking was also revealed, $F(1, 15) = 5.41$, $p = .04$, $\eta_p^2 = .27$. Finally, there was a non-significant SOA x mask presence interaction, $F(2, 30) = 2.98$, $p = .07$, $\eta_p^2 = .17$. T2 response times were not analyzed because the task did not require a speeded response.

Experiment 5b (T2 Prioritized) Results

Memory Task (T1). A repeated measures ANOVA on WM performance revealed no main effects of SOA and the presence of a second task (*p*s > .37, η_p^2s < 0.05), nor an interaction of SOA and presence of a second task, $F(2, 30) = .579$, $p = .57$, $\eta_p^2 = 0.04$. There was a significant effect of mask presence, $F(2, 30) = 42.58$, $p < .001$, $\eta_p^2 = 0.74$.

Parity Judgment Task (T2). There was an average T2 performance of 92.2% (SD = .11). The analysis of T2 response accuracy revealed no significant effects of SOA, or mask presence (both *p*s > .47). There was not a significant interaction between SOA and the presence of a mask, $F(2, 30) = .429$, $p = .66$, $\eta_p^2 = 0.03$. T2 response times were not analyzed because the task required a deferred response.

Between Subjects Analysis (Experiments 5a and 5b). When analyzing Experiments 5a and 5b together (WM accuracy), using instructed priority as a between-subjects factor, there was no significant SOA x presence of a second task x instructed priority interaction, $p = .69$, BF = .117. There was a significant main effect of instructed priority, $p = 0.04$. See Table 5 for full results.

Table 5

Experiment 5 (WM accuracy) repeated measures ANOVA with instructed priority as a between-subject factor

	F	*df*	*P*	η_p^2	*BF*
SOA	.262	2,60	.771	.009	.035
Presence of a second task	1.937	1,30	.174	.061	.543
Mask	123.64	1,30	<.001	.805	$2.59e^{38}$
Instructed Priority	4.175	1,30	.038	.136	1.710
SOA x Instructed Priority	.003	2,60	.968	.001	.063
Presence of a second task x Instructed Priority	.103	1,30	.751	.003	.161
Mask x Instructed Priority	.203	1,30	.656	.007	.181
SOA x Presence of a second task	2.037	2,60	.139	.064	.188
SOA x Mask	2.446	2,60	.095	.075	.426
Presence of a second task x Mask	.222	1,30	.641	.007	.163
SOA x Presence of a second task x Instructed Priority	.378	2,60	.687	.012	.117
SOA x Mask x Instructed Priority	.092	2,60	.912	.003	.093
Presence of a second task x Mask x Instructed Priority	.046	1,30	.832	.002	.207
SOA x Presence of a second task x Mask	1.963	2,60	.149	.061	.474
SOA x Presence of a second task x Mask x Instructed Priority	.953	2,60	.391	.031	.406

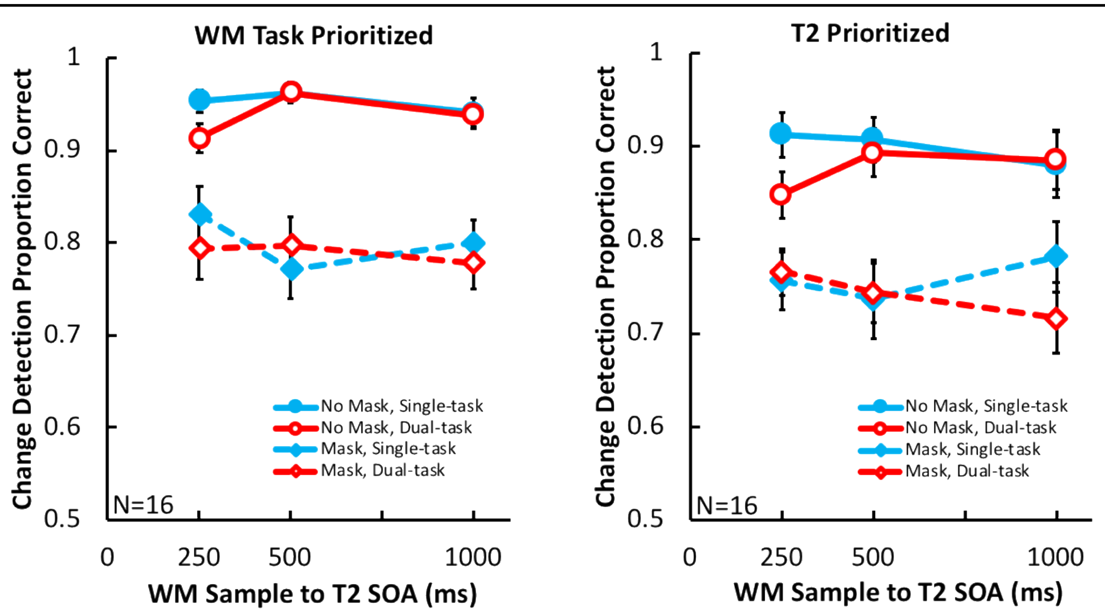

Figure 7. (Left) Performance on WM task when WM was given higher priority than the parity judgement task. (Right) Performance on WM task when parity judgement was given higher priority than WM task. Error bars show the standard error of the mean.

Discussion

A significant main effect of priority provides support that the arbitrary points and timeouts given based on performance indeed manipulated priority, because overall WM performance was greater when the WM task was assigned high priority than when the parity task was assigned high priority. As hypothesized, when the WM task (T1) was prioritized, prolonged consolidation was not obtained. However, contrary to the hypothesis, this was also true when the parity task (T2) was prioritized as well. It was expected that the T2 prioritization condition (Experiment 5b) would lead to the RI effect, similar to Experiments 1 and 2 and to the results of Nieuwenstein and Wyble (2014). Instead, the RI effect was abolished in both experiments. As there are numerous possible explanations for these results, we initially sought to understand if the prioritization account is viable at all. Thus, Experiment 6 tested this same account (including explicit performance feedback), but with an immediate response to T2.

EXPERIMENT 6

The results found in Experiment 5 led to the belief that the deferred parity judgment response could abolish the SOA x presence of T2 interaction for unknown reasons. Thus, it was decided to re-test explicit prioritization in a new experiment modeled after the original change detection manipulation (Experiment 2), where there was an immediate (speeded) parity response. Experiment 6a matched Experiment 5a in terms of the assigned strength of priority to T1 and T2. Experiment 6b matched Experiment 5b. The hypothesis was that the interaction of SOA and presence of a second task would be abolished in Experiment 6a (WM prioritized), but not Experiment 6b (T2 prioritized).

Method

Participants. Data from 16 undergraduate students (Experiment 6a: 4 males, 19-27 years old, M = 21.6 years, SD = 2.75; Experiment 6b: 2 males, 18-33 years old, M = 21.7 SD = 4.44) were collected for Experiment 6a and 6b. Participants were compensated course credit for their completion of the task.

Design and procedure. Experiments 6a and 6b followed a combination of Experiment 2 and Experiments 5a and 5b. Tasks 1 and 2, and the temporal order of responses matched Experiment 2, but also included the added manipulation of feedback that was used in Experiments 5a and 5b.

Experiment 6a Results

Memory Task (T1). A repeated measures ANOVA of WM performance revealed no main effect of SOA ($F(2, 30) = 2.50$, $p = .10$, $\eta_p^2 = 0.14$). There was a significant interaction of SOA and presence of a second task, $F(2, 30) = 4.29$, $p = .02$, $\eta_p^2 = 0.22$. There were also significant main effects of the presence of a second task and mask (both ps < .001).

Parity Judgment Task (T2). There was an average T2 performance of 96% (SD = .06; see Table 2 for a full list of T2 mean accuracy across conditions). The analysis of T2 response accuracy revealed a non-significant effect of SOA, $F(2, 30) = .569$, $p = .572$, $\eta_p^2 = 0.037$. There was a significant main effect of the presence of a mask, $F(1, 15) = 7.29$, $p = .016$, $\eta_p^2 = .38$). The analysis revealed a non-significant interaction of SOA by presence of a mask, $F(2,30) = 2.23$, $p = .125$, $\eta_p^2 = .13$.

T2 response times were also analyzed and revealed a significant main effect of the presence of a SOA, $F(2, 30) = 10.08$, $p < .001$, $\eta_p^2 = 0.40$. The main effect of masking and interaction between SOA and the presence of a mask were not significant (both *p*s > .13). See Table 2 for a full list of mean T2 RTs across conditions.

Experiment 6b Results

Memory Task (T1). A repeated measures ANOVA of WM performance revealed no main effect of SOA ($F(2, 30) = 1.11$, $p = .34$, $\eta_p^2 = 0.07$). There was a significant interaction of SOA and presence of a second task, $F(2, 30) = 4.69$, $p = .02$, $\eta_p^2 = 0.24$. There were significant main effects of the presence of a second task and presence of a mask (both *p*s < .001).

Parity Judgment Task (T2). There was an average T2 performance of 97.9% (SD = .04; see Table 1 for mean accuracies). The analysis of T2 response accuracy revealed a non-significant effect of SOA ($F(2, 30) = 0.147$, $p = .864$, $\eta_p^2 = 0.01$), and a significant effect of masking, $F(1, 15) = 0.83$, $p = .038$, $\eta_p^2 = 0.05$. The analysis revealed a marginally significant interaction of SOA by presence of a mask, $F(2,30) = 2.88$, $p = .07$, $\eta_p^2 = 0.16$.

T2 response times were also analyzed and revealed a significant main effect of SOA, $F(2, 30) = 5.47$, $p = .009$, $\eta_p^2 = 0.27$. No significant main effect of a mask presence, nor interaction of SOA by presence of a mask were revealed (both *p*s > .64).

Between-Subjects Analysis (Experiment 6). When analyzing WM accuracy from Experiments 6a and 6b together with instructed priority as a between-subjects factor there was no significant SOA x presence of a second task x instructed priority interaction, $p = .88$, BF = .096 (~10:1 evidence against the inclusion of the interaction; see Table 6 for all results).

Table 6
Experiment 6 (WM accuracy) repeated measures ANOVA with instructed priority as a between-subject factor

	F	*df*	*P*	η_p^2	*BF*
SOA	3.005	2,60	.057	.091	.260
Presence of a second task	90.092	1,30	<.001	.750	$4.52e^{22}$
Mask	207.857	1,30	<.001	.874	$1.17e^{43}$
Instructed Priority	.111	1,30	.741	.004	.311
SOA x Instructed Priority	.003	2,60	.997	.000	.058
Presence of a second task x Instructed Priority	.401	1,30	.531	.013	.205
Mask x Instructed Priority	.022	1,30	.883	.001	.155
SOA x Presence of a second task	8.745	2,60	<.001	.226	145.49
SOA x Mask	4.509	2,60	.015	.131	1.029
Presence of a second task x Mask	3.239	1,30	.082	.097	.751
SOA x Presence of a second task x Instructed Priority	.128	2,60	.880	.004	.096
SOA x Mask x Instructed Priority	.334	2,60	.717	.011	.121
Presence of a second task x Mask x Instructed Priority	.769	1,30	.388	.025	.316
SOA x Presence of a second task x Mask	1.529	2,60	.225	.048	.298
SOA x Presence of a second task x Mask x Instructed Priority	.103	2,60	.903	.003	.238

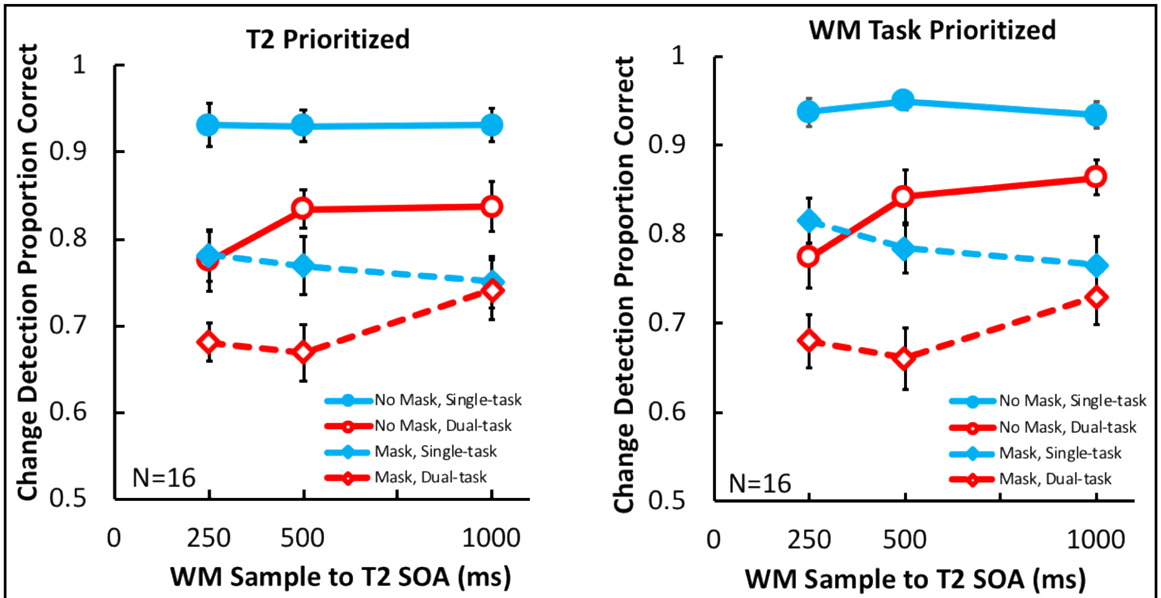

Figure 8. (Left) Performance on the WM task when WM task was given higher priority than the parity task. (Right) Performance on WM task when parity task was given higher priority than the WM task. Error bars show the standard error of the mean.

Discussion

Regardless of task priority, identical RI effects—up to one second were observed, eliminating the account that RI with WM consolidation results from strategic prioritization of T2 (the 2AFC parity task) over WM encoding. This result supports the structural account of slow consolidation because the pace of consolidation measured via RI was not affected by explicit task prioritization. If strategic factors do not affect the pace of WM consolidation, it further suggests that Nieuwenstein and Wyble (2014)'s RI approach can no longer be thought of as prolonging an otherwise-rapid consolidation process (c.f., Gegenfurtner & Sperling, 1993; Vogel et al., 2006; Woodman & Vogel, 2005). Instead, these results support the account that consolidation is always slow, and that prior measurement approaches were insensitive to unsuccessful consolidation. It should be noted that Nieuwenstein and Wyble (2014) formulate a similar argument, suggesting that the key assumption of most prior studies of consolidation rate—that visual masks disrupt consolidation—is erroneous. With this shift in interpretation, Experiments 7 and 8 turn to characterizing the limitations of the RI technique for measuring consolidation rate. More specifically, Experiments 7 and 8 investigate boundary conditions for RI on consolidation by a second task. Because a deferred response to T2 abolished the RI found in the previous experiments (c.f. Experiments 4 and 5), there may be a critical role of the immediate T2 response for driving RI with consolidation.

EXPERIMENT 7

While RI on WM consolidation was never obtained when the T2 (parity decision) response was deferred until after the WM response, the RI effect was constantly shown when an immediate T2 response was required. This could be due to the immediate, speeded T2 response, being a necessary condition of RI on WM consolidation. One counterexample in the literature was found where the RI effect was obtained with an unspeeded T2 response (Ouimet & Jolicoeur, 2007). However, the arrangement of Ouimet and Jolicoeur's task was fundamentally different than the tasks used in the current study because their T2 required a second WM sample rather than a T2 decision task. A second mask was also included after the presentation of the second memory sample. Ouimet and Jolicoeur's results suggest that RI can be obtained with no task switch, and a deferred response execution. Due to the conflicting results, Experiment 7 incorporates these task differences to test whether RI on WM consolidation can be obtained with a deferred T2 response.

Method

Participants. Data from 16 undergraduate students (3 males; 18-38 years old, mean (M) age = 22.89 years, SD = 5.21) were collected. Participants were compensated course credit for their completion of the task.

Design and procedure. A within-subject design was used for Experiment 7, with SOA, the presence of a mask, the presence of a second task, and the presence of a second mask as factors. The SOA was either 150, 250, 500, or 1,000 ms. The additional 150 ms SOA was included because after close inspection of the results from Experiments 4 and 5 (Figures 6 and 7), at least in the unmasked conditions, there could have been diminished WM performance on T2-present trials at the shortest (there, 250 ms) SOA. Findings at the shortest SOA from Experiments 4 and 5 could have reflected a more rapid consolidation process that was still susceptible to RI from T2, but only in an earlier time window than evaluated in Experiments 4 and 5. To allow for this additional, shorter SOA, the duration of the masks decreased from 100 ms to 50 ms.

Instead of a displaying a digit and having participants judge whether it was odd or even, participants had to simply remember the presented digit for later recall. This effectively switched T2 from a parity decision task to a second WM encoding task. All trial types were randomly intermixed, leading to 512 total trials (32 conditions; 16 trials per condition).

Figure 9 shows a visual of the design used in Experiment 7. Each trial began with four place holder crosses located in the center of the screen. Participants initiated the trial by using the spacebar. The letters then took the place of the crosses and were shown for 100 ms. Half the trials presented four "#" symbols for 50 ms to mask the WM sample. If a visual mask was present, participants were told to ignore the mask. For the other half of the trials that did not include the mask, there was a blank screen for a 50 ms period. This was then followed by a 0, 100, 350, or 850 ms blank screen interval. This blank interval was followed by an additional 100 ms blank screen (for single-task trials) or the display of a single digit WM sample for 100 ms (for dual-task trials). For both single-task trials (after blank interval) and dual-task trials (after digit presentation) a second mask ("#") would appear on half of the trials for 50 ms. Again, if a second visual mask was present, participants were told to ignore the mask. There was then a final 950 ms blank interval. Participants were then shown a new array of four letters that remained on the screen until a response was made. They used one of two keys to respond to the change detection portion of the task. After a response was made, if it was a dual-task trial, they were asked, "which digit did you see?" and had to input the digit using the corresponding key on the keyboard.

Results

As for the previous experiments, analysis focused on T1 WM report accuracy. A repeated measures ANOVA was conducted with SOA, presence of mask 1, presence of mask 2, and presence of a second task as factors. There were no significant SOA x presence of a second task [x mask1 and/or mask2] interactions (all ps > 0.16, all η_p^2s < 0.10, all BFs < .180). (See Table 7).

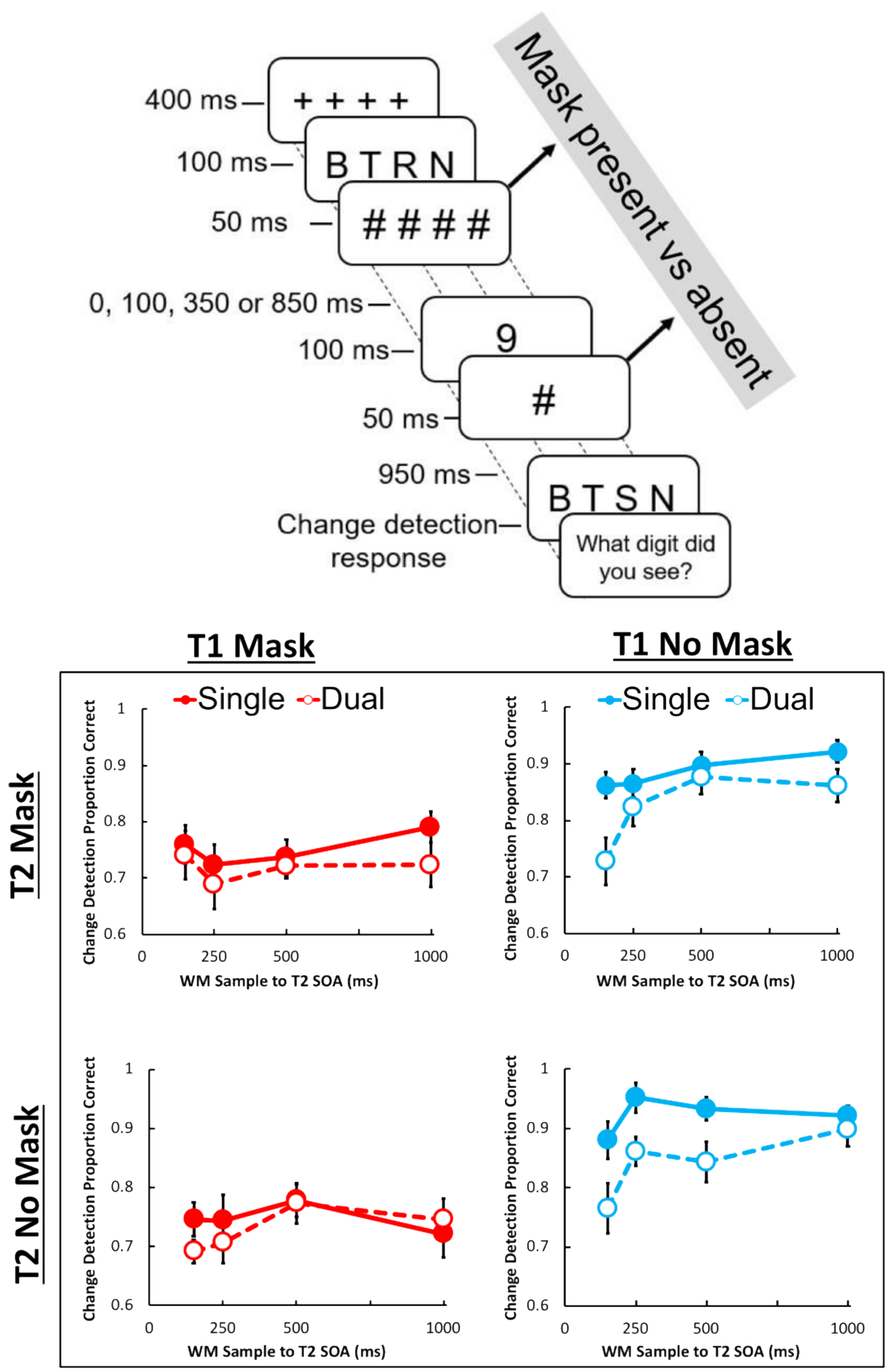

Figure 9. Task design for dual-task trials shown on the top. Results for WM performance are shown on the bottom in a two-by-two grid.

Table 7
Experiment 7 (WM accuracy) repeated measures ANOVA

	F	df	P	η_p^2	BF
SOA	10.709	3,45	<.001	.417	36.127
Presence of a second task	36.955	1,15	<.001	.711	25146.83
Mask 1	125.934	1,15	<.001	.894	8.023×10^{32}
Mask 2	1.744	1,15	.206	.104	.330
SOA x Presence of a second task	1.484	3,45	.232	.090	.124
SOA x Mask 1	3.537	1,15	.022	.191	4.491
SOA x Mask 2	1.025	1,15	.391	.064	.087
Presence of a second task x Mask 1	6.676	3,45	.021	.308	2.155
Presence of a second task x Mask 2	.0002	3,45	.989	.000	.134
Mask 1 x Mask 2	2.145	1,15	.164	.125	.313
SOA x Presence of a second task x Mask 1	.081	3,45	.495	.051	.084
SOA x Presence of a second task x Mask 2	1.810	3,45	.159	.108	.180
SOA x Mask 1 x Mask 2	1.466	1,15	.236	.089	.239
Presence of a second task x Mask 1 x Mask 2	.008	3,45	.406	.046	.221
SOA x Presence of a second task x Mask 1 x Mask 2	.467	3,45	.707	.030	.163

Shortest SOA results. Upon examining the results shown in Figure 9, there appears to be a dual-task effect at the shortest SOA. This same effect does not seem present in the longer SOAs. To further investigate whether there were dual-task effects at the shortest SOA a repeated-measures ANOVA only using the data from the 150 ms SOA, with the presence of a second task, presence of mask 1, and the presence of mask 2 as factors. Significant main effects of the presence of a second task and mask 1 were obtained (both ps < .003, η_p^2 > .47, BFs > 91.53). There was a significant presence of a second task x presence of mask 1 interaction, $F(1,15) = 6.345$, $p = 0.02$, $\eta_p^2 = .30$, BF = 2.17. All other main effects and interactions were non-significant (all ps > .11, η_p^2s < .17, BFs < .68). A simple main effects analysis revealed a significant difference between single- and dual-task trials in the unmasked (mask 1) condition, $F(1,15) = 20.170$, $p <$

.001, η_p^2 = .57, BF = 99.76. There was not a significant difference between single- and dual-task trials in the masked 1 condition, $F(1,15)$ = 1.731, p = .21, η_p^2 = .10, BF = 1.39.

Discussion

There was one instance found in the literature in which researchers obtained RI in a dual-task paradigm similar to that of the present investigation and to that of Nieuwenstein and Wyble (2014) even though T2 report was delayed until the end of the trial, after T1 report (Ouimet & Jolicoeur, 2007). These results oppose and challenge the present findings. One key difference between the current study and Ouimet and Jolicoeur (2007)'s study is that instead of having participants complete a decision-making task (parity judgment) as T2, they simply had participants complete a second WM task, where participants had to encode another array of digits for later recall. This suggests that the RI effect was obtained even without a task switch, and with delayed response (motor) execution—a surprising result given the human ability to successfully encode serially-presented WM items (Di Lollo, Kawahara, Shahab Ghorashi, & Enns, 2005; Olivers, van der Stigchel, & Hulleman, 2007; Potter, Chun, Banks, & Muckenhoupt, 1998). Experiment 7 was unable to replicate Ouimet and Jolicoeur's (2007) results. The current experiment further confirms that an immediate T2 response is necessary to obtain RI with WM consolidation. The finding from Ouimet and Jolicoeur, 2007 may be a potential outlier; this view is further supported by additional experimentation within their own study in which they also did not obtain the RI effect in very similar manipulations. It must also be noted that their study was underpowered, and only included results from a total of ten participants.

Experiment 7 also included the shortest T1-T2 SOA yet tested. This decision was driven by the observation that the implicit prioritization experiments (Experiments 4 and 5) seemed to yield poorer WM accuracy in the dual-task condition at the 250 ms SOA than at later SOAs, even though those experiments had no statistically significant SOA x presence of a second task interaction (Figure 7). There also seemed to be a difference between single- and dual-task trials at that shortest SOA. Taken together, the 250 ms SOA results of Experiments 4 and 5 raised the possibility that there was a more rapid consolidation process that was still vulnerable to RI from a second task, but only at an earlier time window. Thus, Experiment 7 included a shorter SOA of 150 ms to provide a further test of the strategic account of slow consolidation. Although a difference between single- and dual-task trials at the shortest SOA was observed in Experiment 7, this effect only occurred when there was no T1 mask. That the presence or absence of a T1 mask modulated the effect is critical because it suggests that the effect is not driven by T2 retroactively interfering with WM consolidation at the earlier time window. Instead, it is consistent with T2 at very short SOAs (Experiments 4 and 5: 250 ms, Experiment 7: 150 ms) acting as a mask. This would parsimoniously explain both why the results for no-mask-1 dual-task trials at the shortest SOA look very similar to that of the single- and dual-task trials in the masked conditions, and more broadly, why there is an (apparent) dual-task effect at short SOAs that does not persist to longer SOAs in Experiments 4, 5, and 7.

EXPERIMENT 8

Results from Experiments 4 and 5 (deferred T2 response) suggest that an immediate T2 response may be necessary to disrupt ongoing WM consolidation.

However, it is still unclear on whether an intervening decision task with an immediate T2 response during the WM interval is necessary to obtain RI with WM consolidation, or if it is possible to obtain the RI effect with sensory decision during the WM interval that does not require a response. The goal of Experiment 8 was to determine if making a sensory decision during the WM delay interval is sufficient to drive RI with WM consolidation. This experiment matched the setup of Experiment 2, but instead of only including single-task and dual-task trials, single-task-with-distractor trials were added and randomly intermixed. This arrangement would determine if a distractor without a motor response was sufficient to drive RI with WM consolidation by making comparisons between single-task versus single-task-with-distractor WM accuracy, as well as single-task versus dual-task WM accuracy (Experiment 2 replication). Importantly, because dual-task and single-task-with-distractor trials were randomly intermixed, single-task-with-distractor trials still required a sensory decision in order to determine that the presented stimulus was a distractor and not a T2.

Methods

Participants. Data from 16 undergraduate students (3 males; 18-49 years old, M = 24.1 years, SD = 7.76) were included in the analysis for Experiment 8. They were compensated course credit for their time.

Materials, design, and procedure. Before beginning Experiment 8, participants completed a k-estimation task (see Experiment 3 *Method*, above). Participants were given a short break before starting Experiment 8.

Experiment 8 was similar to Experiment 2 with the exception of different SOA timing (i.e. excluding the 500 ms SOA and adding a 150 ms SOA), and mask timing (50 ms masks instead of the original 100 ms). For Experiment 8 a condition was added in which some 2AFC targets were replaced with distractors (while actual 2AFC targets were still presented on randomly-intermixed trials). The distractor was a “%” symbol. Thus, distractor trials required classification of the delay-period stimuli as distractors and not targets, but required no motor response. Experiment 8 was a 3x3x2 within-subjects design with SOA (150, 250, 1000 ms), presence of second task or distractor (single-task, dual-task, single-task-with-distractor), and presence of a mask as factors (18 condition types at 20 trials each; 360 total trials). All condition types were randomly intermixed.

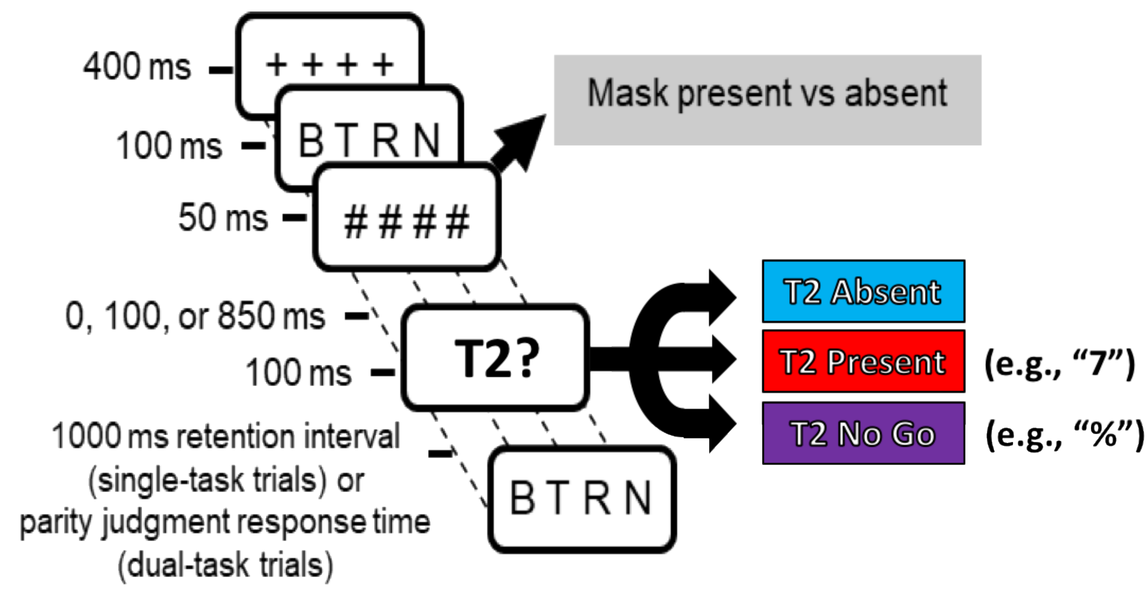

Figure 10. Task design for Experiment 8. Note: In a third of the trials each participant was shown a blank screen (single-task, no response needed), a digit (parity response needed), and percent sign (single-task-with-distractor, no response need). This is an example of a “no change” trial.

Results

Memory Task (T1). WM performance was examined with a repeated-measures ANOVA including SOA, the presence of a mask, and 3-level T2 (single-task, single-task-with-distractor, and presence of parity judgment) as factors. There was a marginally significant main effect of SOA, $F(2,30) = 2.62$, $p = 0.09$, $\eta_p^2 = 0.15$. The analysis also revealed significant main effects of the presence of a second task, presence of a mask, and SOA by T2 interaction (all ps < 0.01 and $\eta_p^2 s > 0.22$). No other interactions were statistically significant (ps > .10).

Results excluding dual-task trials. WM accuracy (change detection) was measured using a repeated-measures ANOVA with SOA (150, 250, and 1000 ms), presence of a distractor (single-task and distractor conditions; excluding dual-task), and presence of a mask as factors. There were no significant main effects of SOA and presence of a distractor (both ps > 0.36 and $\eta_p^2 s < 0.06$). There was no significant interaction of SOA by the presence of a distractor, $F(2,30) = 2.70$, $p = 0.08$, $\eta_p^2 = 0.15$. A non-significant interaction suggested that a distractor (i.e., a sensory decision without an accompanying motor response) is insufficient to evoke RI on WM consolidation.

Results excluding single-task-with-distractor trials. To verify that the normal RI effect was obtained within this experiment, another analysis was completed which included the presence of a second task as a factor instead of the single-task-with-distractor condition. There were significant main effects of the presence of a second task, and presence of a mask (both ps < 0.001, $\eta_p^2 s > 0.78$). There was a significant interaction of SOA by the presence of a second task, $F(2,30) = 7.40$, $p = 0.002$, $\eta_p^2 = 0.33$. The significant interaction replicated the earlier findings of RI by a 2AFC task on WM consolidation (Experiments 1, 2, and 6).

Parity Judgment Task (T2). There was an average T2 performance of 88% (SD = .12; see Table 1 for mean T2 performance). The analysis of T2 response accuracy revealed no significant main effects or interaction between SOA and the presence of a mask (all *p*s > .19 and η_p^2s < 0.10).

T2 response times were also analyzed and revealed a significant main effect of masking, $F(1, 15) = 7.65$, $p = .014$, $\eta_p^2 = 0.34$. There was no significant main effect of SOA, nor an interaction of SOA by presence of a mask (both *p*s > .39 and η_p^2s < 0.06).

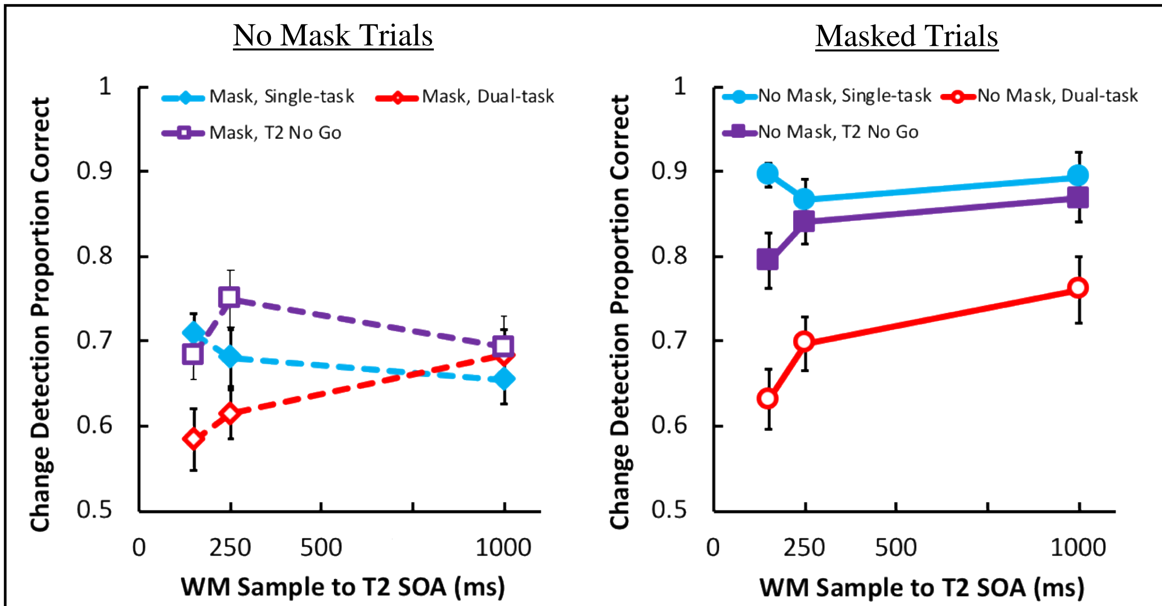

Figure 11. WM accuracy for unmasked trials (left) and masked trials (right). Error bars show the standard error of the mean.

Discussion

In Experiment 8, a condition was added where 2AFC targets were replaced with distractors (while actual 2AFC targets were still presented on randomly-intermixed trials). Thus, distractor trials required classification of the delay-period stimuli as distractors and not targets, but required no motor response. On these distractor trials, similar results to trials with no delay-period task or stimulus were obtained. This result rules out the idea that distraction during the WM delay by the mere presence of a transient visual stimulus, or even by non-motoric cognitive processes such as classifying stimuli as targets vs. distractors, is sufficient to drive RI. Instead, a motor response appears necessary for RI. Experiment 9 further investigates whether RI is obtained due to response selection (with a motor response), response execution, and/or task switching between the 2AFC and WM tasks.

EXPERIMENT 9

It is still an open question as to what aspects of an intervening second task causes disruption of ongoing WM consolidation. Is a mere speeded motor response with little or no decision-making demand sufficient to interfere with WM consolidation? The goal of Experiment 9 was to further extend the findings from Experiment 8. Results from Experiment 8 suggested that a sensorimotor decision and response—rather than the mere distraction from presentation of a stimulus during the WM delay interval, or even from a sensory decision without a motor response during the WM delay interval—were necessary to obtain RI. However, Experiment 8 left open the question, is consolidation disrupted by motor response, or by mapping of sensory motor decision to a motor response? In Experiment 8 when the distractor appeared (e.g., "%") participants were instructed to not make a response, and continue on to the WM change detection response when presented at the end of the trial. However, Experiment 9 required participants to make a T2 response, but for a less demanding T2 task than the previously used parity judgment. The experimental design of Experiment 9 matched Experiment 2 with one key difference: instead of T2 requiring a speeded 2AFC parity judgment, participants now provided a speeded response using only a single key to detect whether or not a digit was present. If a digit was presented during the WM delay interval, they were required to make a speeded response using a single designated key regardless of whether the digit was odd or even. Using such a simple reaction time/stimulus detection task allowed for the determination of whether a mere speeded motor response with little or no decision-making demand is sufficient to interfere with WM consolidation. Along with this, the less demanding T2 allowed for the examination of the significance (or insignificance) of response selection in driving RI with WM consolidation. Specifically, Experiment 9 echoes a question raised by Nieuwenstein and Wyble (2014, their Experiment 3) which used a color discrimination (2AFC) vs color detection task (similar to Experiment 9) as T2. Unfortunately, they only reported results for the comparison of the two tasks, omitting whether detection was sufficient to obtain RI with WM consolidation.

Method

Participants. Data from 16 undergraduate students (6 males; 18-33 years old, M = 21.3 years, SD = 3.96) were collected for the analysis. One participant was excluded from the analysis due to a difference in SOAs used. Data from an additional participant was unable to be acquired due to COVID-19. Participants were compensated course credit for their time.

Materials, design, and procedure. Experiment 9 used the same setup as Experiment 2. However, in Experiment 9 instead of an intervening 2-AFC parity judgment task, participants completed a stimulus detection task. If a digit was presented, participants were instructed to make a speeded response using the comma key, regardless of what the digit was.

Experiment 9 consisted of a 3x2x2 within-subject design with SOA (250, 500, and 1000 ms), presence of a mask, and presence of a second task as factors. This equates to 12 conditions with 26 trials for each condition (total of 312 trials).

Results

Memory Task (T1). A repeated-measures analysis of variance (ANOVA) was used to examine the change detection performance including SOA, the presence of a mask, and the presence of a second task (stimulus detection) as factors. There was a significant main effect of SOA, $F(2,28) = 5.55$, $p = 0.009$, $\eta_p^2 = 0.28$. There were significant main effects revealed for the presence of a mask, and presence of a second task (both ps < .001 and η_p^2s > .62). However, the interaction of SOA with the presence of the second (detection) task was very small ($\eta_p^2 = 0.03$) and did not achieve statistical significance ($p = .65$, BF = .214).

Parity Judgment Task (T2). There was an average T2 performance of 96% (SD = .06; see Table 1 for mean T2 accuracy across conditions). The analysis of T2 response accuracy revealed a significant effect of masking $F(1, 14) = 8.26$, $p = .012$, $\eta_p^2 = 0.37$. There were no significant main effects of SOA, nor an interaction between SOA and mask presence (both ps > .48 and η_p^2s < 0.05).

T2 response times were also analyzed and revealed no significant effects of SOA, mask, or interaction of SOA by mask (all ps > 0.32 and η_p^2s < 0.08). See Table 1 for a list of mean T2 RTs across conditions.

Discussion

It is still an open question about what aspects of T2 causes disruption of ongoing VWM consolidation. Experiment 9 tested whether a motor response with little or no decision-making demand was sufficient to interfere with WM consolidation. Instead of having participants complete an intervening 2AFC parity judgment decision task, in Experiment 9 participants completed a simple stimulus detection task that required a response using a single key if a digit appeared during the WM delay. By having participants complete a stimulus detection task it was determined that a mere speeded motor response is not sufficient to interfere with WM consolidation. Considering Experiments 8 and 9 together, it appears that RI is obtained due to the mapping of a sensory motor decision to a motor response.

ARE RI AND PI MUTUALLY EXCLUSIVE?

Thus far the present study has been concerned with the speed of WM consolidation, which can be measured by evaluating WM accuracy across varying SOAs. However, the study has yet to capitalize on the fact that T2 RTs have been collected. Analyzing these data can provide additional insight into the current results. RI is marked by T2 influencing T1 performance. However, there is another well-studied phenomenon known as the psychological refractory period (PRP) effect which is when T2 suffers from PI by T1 (Pashler, 1994; Telford, 1931; Welford, 1952). The PRP effect can be observed when looking at changes in T2 response time as a function of T1-T2 SOA. When there is a shorter SOA from T1 to T2, response times for T2 tend to be slower, recovering as SOA increases (Pashler, 1994; Telford, 1931; Welford, 1952). This is typically interpreted as T1 using a resource or bottlenecking a process that is also needed for T2; the "central" processing time for T1 can thus be inferred from the SOA at which T2 responses return to their usual timing (Marois & Ivanoff, 2005; Pashler, 1994). A PRP effect is detected by a main effect of SOA with longer response times at shorter SOAs. Nieuwenstein and Wyble (2014) obtained RI, but failed to observe consistent PI using this retroactive interference approach. However, Jolicoeur and Dell'Acqua (1998) obtained PI, but not RI, using a similar paradigm. Because of these conflicting findings, there is an open question on whether RI and PI can be observed together using the same task. Pooling together and analyzing the T2 RT data would speak to whether or not something about the present paradigm is leading to PI vs. RI. The goal of the current analysis is to see if RI and PRP (PI) effects can be obtained at the same time. Data across six experiments have been pooled together for the current analysis. Participants were excluded from experiments where T2 response was deferred (i.e. Experiments 3, 4, and 6), and where the WM sample differed (i.e. Experiment 3: Color patches); thus, in this analysis, data from Experiments 1, 2, 6a, 6b, 8, and 9 were used.

Method

Participants. Data from 95 undergraduate students (24 males; 18-49 years old, M = 22.6 years, SD = 5.97) were included in the analysis.

Materials, design, and procedure. A mixed ANOVA on T2 RTs with SOA and mask as within-subject factors, and experiment as a between subject factors were used.

Because all of the experiments used in this analysis included both a 250 ms and 1000 ms SOA, these SOAs were used to define RI and PI.

Results

Response times for T2 were analyzed. There was a significant main effect of SOA on RTs, $F(1, 89) = 28.14$, $p < .001$, $\eta_p^2 = 0.24$. There was no significant interaction between SOA and experiment, $F(5, 89) = 2.08$, $p = .08$, $\eta_p^2 = 0.10$. There was a significant main effect of experiment, $F(5, 89) = 8.14$, $p < .001$, $\eta_p^2 = 0.31$. All other effects and interactions were not significant (all ps > .36 and η_p^2s < 0.10).

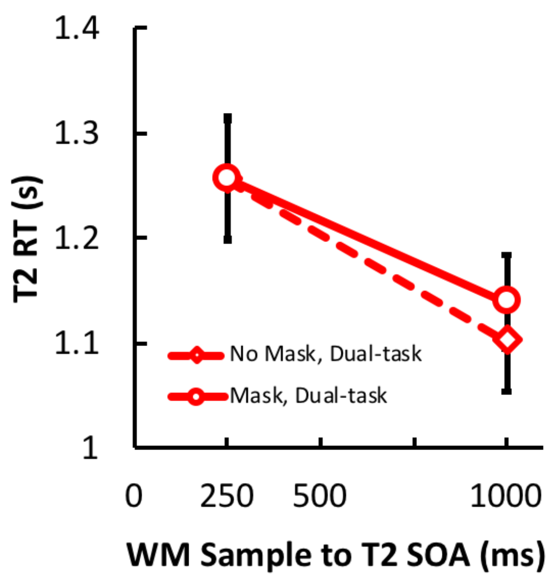

Figure 12. Graph includes T2 RT in seconds across SOA. Error bars show the standard error of the mean.

Discussion

A significant main effect of SOA revealed from a mixed-ANOVA on T2 RT data suggest there was PI. This supports the hypothesis that both RI and PI can be observed together using the retroactive interference paradigm. The current results suggest that each of the previous studies (Nieuwenstein & Wyble, 2014; Jolicoeur & Dell'Acqua, 1998) may not have had the sufficient power to observe one or the other effect, but it is not the case that it is impossible to observe both effects together.

IS THERE A RELATIONSHIP BETWEEN RI AND PI?

Since both RI and PI were successfully observed together in the same task, the goal of the current analysis involved investigating whether there is a relationship between RI and PI. For instance, if greater PI was observed when lesser RI was obtained (and vice versa), it would suggest a tradeoff in central resource allocation between the two tasks. If there was no relationship, it would suggest that there is not a tradeoff in central resource allocation between the two tasks. The current analysis examined whether there is an interference tradeoff at both the subject- and trial-level.

First, it was examined whether there was a relationship between RI and PI on the participant-level. Examining whether there is a relationship between RI and PI would provide insight on whether or not there are trait-level propensities. Furthermore, if there was a relationship, it would suggest there was a trait-level predisposition towards RI or PI. RI and PI were calculated for each participant, and correlated using a Pearson's r. Based on RI results from experiment 6 it was hypothesized that no such tradeoff would be obtained given that explicit prioritization of T1 does not seem to diminish RI of T2 on T1.

Examining across-subject data for a relationship between RI and PI provides insight about whether there are trait-level propensities for the task (e.g., individual differences in skills at the two tasks). However, determining whether engaging in a strategy that reduces RI, trades off with PI (in other words leads to greater PI) and vice versa could provide additional insight about WM consolidation. Therefore, tradeoffs at the trial-level were examined by comparing WM accuracy across SOAs on trials with high vs. low PI. To do so, a median split of T2 RTs was used, and examined whether there was a difference in WM accuracy.

Method

Participant-level. Here, both RI and PI were related in 95 participants drawn from six dual-task experiments (Experiments 1, 2, 6a, 6b, 8, and 9). RI was operationalized as the difference in WM report accuracy for dual-task 1000 ms SOA minus 250 ms SOA for each participant. PI was operationalized as the difference in T2 RTs for 250 ms SOA minus 1000 ms SOA for each participant. Only common SOAs across all experiments (250 ms & 1000 ms) were used.

Trial-level. A median split within-participant, within-condition of T2 RT was used. A mixed ANOVA on WM accuracy with SOA (250 ms and 1000 ms), presence of a mask, and T2 RT split as within-subject factors, and experiment as between subject factors was used.

Results

Participant-level. There were no significant correlations between RI and PI for masked (r = .133, p = .199; see Figure 13a) or unmasked conditions (r = -0.085, p = .414; see Figure 13b).

Trial-level. There were significant main effects of SOA, presence of a mask, and experiment (all ps < 0.009, all η_p^2s > 0.16). There was a significant interaction between SOA and T2 RT split, $F(1, 89) = 4.365$, $p = .04$, $\eta_p^2 = 0.05$. All other main effects and interactions were not significant (all ps > .20 and η_p^2s < 0.07). Thus, there was greater RI on the trials with the fastest T2 responses, regardless of mask presence or absence (see figure 13c for WM results).

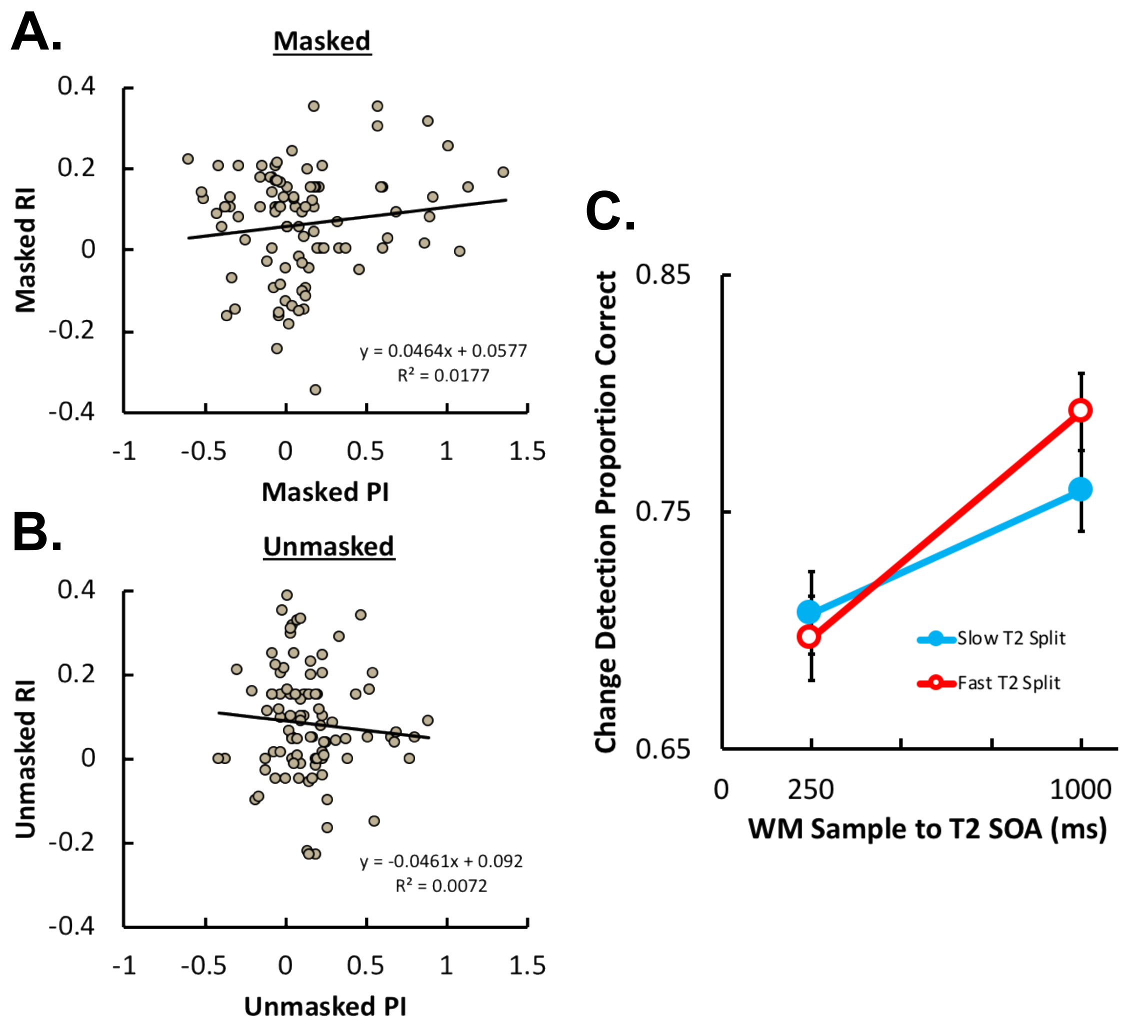

Figure 13. Correlation plots between RI and PI for masked (A) and unmasked (B) conditions. (C) WM accuracy across SOA for slow (high PI trial) and fast T2 (low PI trials) splits. Error bars show the standard error of the mean.

Analysis-level Discussion

A non-significant correlation between RI and PI suggests there is no tradeoff at the participant-level. Furthermore, the across-subjects results suggest that there is not a trait-level predisposition towards PI or RI. On the trial-level, a significant interaction was observed between SOA and T2 RT split. This finding suggests that WM consolidation speed could vary from trial to trial. At the short SOA, there was no difference in WM accuracy for both the low and high PI trials. However, at the long SOA, WM accuracy increased more for the low PI than the high PI trials. This pattern of results can be conceptualized as no difference in WM accuracy at the short SOA for low and high PI trials because WM consolidation was likely not complete regardless since it is estimated to take around 1 second to be complete (see Figure 14A & 14B). However, WM accuracy is higher at the long SOA for low PI trials because WM consolidation must have been closer to completion (as expressed through quicker T2 RTs, and therefore lower PI; see Figure 14C) than high PI trials, where consolidation was still ongoing thus leading to greater PI and slower T2 RTs (see Figure 14D). PI may be caused by delayed parity RS due to ongoing WM consolidation (shown in Figure 14B & 14D as the separation between P2 and parity RS). This idea is consistent with predictions from non-WM-T1 tasks in similar dual task perceptual paradigms (cf. Pashler, 1994). Fast consolidation appears in Figure 13C as a greater slope of the RI function because there was greater WM achieved before T2 onset. A greater slope of the RI function is associated with faster T2 decisions because they reflect faster WM consolidation and thus less PI.

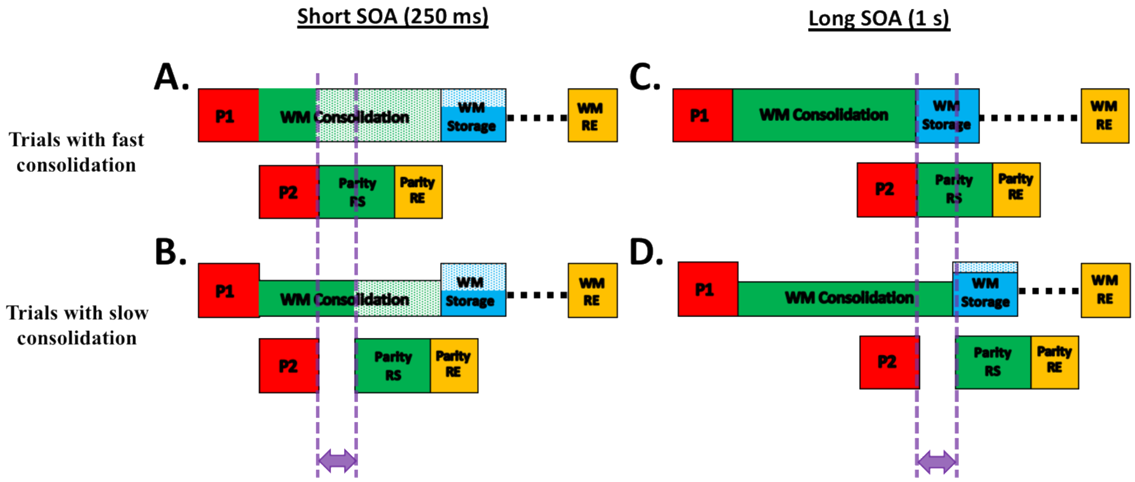

Figure 14. Trial-by-trial variability for WM consolidation speed. Examples of trials with fast WM consolidation (A and C) vs. trials with slow WM consolidation (B and D) across SOA. Abbreviations within boxes are explained in Figure 2. When WM consolidation is fast, it exerts less PI (expressed through T2 response times—time from onset of P2-Parity RE) because it is more likely to be complete before T2 onset. For trials where WM consolidation is slower, there is more PI (parity RS delayed) because consolidation was still ongoing at the onset of T2. Each pair of purple vertical dotted lines represent the change in the duration of WM consolidation. The Height of the WM consolidation bar is shorter for trials with slow consolidation (B and D) because slow consolidation can be explained by reduced resources on that trial.

GENERAL DISCUSSION

The temporal dynamics of WM consolidation are the subject of conflicting results, with masking paradigms suggesting rapid consolidation but the attentional blink suggesting that consolidation may take hundreds of milliseconds. Nieuwenstein and Wyble (2014) showed RI by a speeded decision task during the WM delay on WM consolidation. This effect lasted even longer than the AB—up to at least 1 second. This result presents another conflicting finding on the time course of WM consolidation, and also contradicts the typical dual-task finding of PI by WM consolidation on a subsequent task. The present research sought to understand just what leads to the longer estimate of WM consolidation observed using the retroactive dual-task interference paradigm, with the end goal of determining if prolonged consolidation is a structural (obligatory) feature of WM encoding, or the result of strategic factors. After closely replicating the RI effect on WM consolidation (compare Experiment 1 to Nieuwenstein & Wyble, 2014), Experiment 2 demonstrated that response demands do not explain the results by shifting from whole-report to change-detection paradigms. Specifically, when participants completed randomly intermixed single- and dual-task change detection trials, a significant interaction was observed between delay duration and the presence of a second task, exactly as in whole-report. Next, Experiment 3 revealed that the RI effect is obtained regardless of the type of stimuli used for the WM sample (visual verbal stimuli vs. visual non-verbal stimuli), which is consistent with a common, central consolidation mechanism rather than distinct consolidation mechanisms for different kinds of WM representational formats. Experiments 4-6 both directly and indirectly manipulated task priority to test whether strategic factors mediate RI with WM consolidation. Regardless of task priority, identical RI effects were observed, eliminating the account that slow consolidation results from strategic prioritization of the 2AFC task over WM encoding. The current findings support the notion that the longer estimate observed in RI tasks is a structural feature of WM encoding, rather than being caused by strategic factors (Experiments 4-6). This finding refuted the original hypothesis that strategic factors would influence the speed of WM consolidation, and led to the further investigation of what aspects of a second task are sufficient to drive RI with WM consolidation. The present research suggests that having a second task alone does not always lead to RI. For instance, no RI was obtained when the decision task (second task) was unspeeded (Experiments 4 & 5). Moreover, a sensorimotor decision and speeded response—not mere distraction or response—were required to obtain RI (Experiment 8 & 9).

Although not skeptical of the results from Nieuwenstein and Wyble (2014), initially it was suspected that the long (compared to previous estimates) estimate of consolidation duration revealed by the retroactive dual-task interference paradigm was due to the task-specific strategies that participants were implicitly induced to adopt by the experimental design. That is, it was suspected that having participants provide an immediate response to T2 while delaying T1 report could have resulted in participants adopting strategies to perform the task, such as prioritizing T2 over T1. Thus, this implicit prioritization has potential to evoke a longer, drawn-out consolidation process. However, the fact that explicit manipulations of task priority did not abolish the RI effect weighs strongly against this strategic account. Nevertheless, there is still a large discrepancy between WM consolidation estimates, namely visual masking and PI paradigms leading to quicker estimates and RI paradigms leading to slower estimates.

Visual masking paradigms have relied on the assumption that WM consolidation is a unitary process. That is, if a stimulus can be recalled after being masked, the consolidation process must have been complete. However, Nieuwenstein and Wyble (2014) demonstrated that there may be two separate stages to the consolidation process. A two-stage consolidation model can reconcile the contrasting estimates found by various measurement approaches. For instance, visual masking studies assume they are measuring the consolidation process as a whole, but in reality, may only be measuring stage 1, whereas other measurement approaches may be making conclusion based on the second stage of consolidation. Nieuwenstein and Wyble (2014) describes this first stage as the selection-for-consolidation process. This stage is vulnerable to visual masks. However, once an item is selected, stage 2 of consolidation begins, lasting for a longer duration depending on the stimulus. Ye et al. (2017) proposed a similar two-phase VWM resource allocation model. This model consists of an involuntary and voluntary phase. The first (involuntary) phase is described as a bottom-up, or stimulus-driven phase where a low-resolution representation is created of the visual stimulus. Only once this earlier phase is complete can the second (voluntary) phase begin, in which individuals can reallocate resources to create a high-precision representation of the stimulus. These types of two-stage VWM consolidation models can explain why these various measurement approaches come to different conclusions regarding the estimation of the consolidation rate.

Another disparity stems from differences in obtaining PI vs. RI in similar dual-task paradigms. Studies that typically use PI for encoding a sample array into memory on a subsequent T2 rely on stronger assumptions than studies that use RI of T2 on WM performance. That is, they rely on indirect evidence, assuming that it is T1 consolidation driving interference on T2, rather than ancillary processes related to WM that could include maintenance as well as consolidation. RI studies do not rely on this same assumption. Instead, they rely on the assumption that something about T2 drives interference on WM consolidation. The subject of the interference logically must be WM consolidation, because if it were maintenance, it would be expected that RI would be constant across SOAs. However, it is not completely clear what aspect(s) of T2 drive RI. Insight into the boundary conditions of T2 RI on WM consolidation may be gained by manipulating T2 demands and observing changes (or the absence of changes) in the obtained RI. This thesis provides additional insight into the necessary T2 characteristics to evoke RI.

Analyzing T2 response times provided a complementary test of the idea of strategic trade-offs between T1 consolidation and T2 decision making. For instance, if greater RI (in terms of T1 performance) was obtained when lesser PI (in terms of T2 RTs) was observed (and vice versa), this would suggest a strategic tradeoff. However, because at the subject-level there was no relationship between RI and PI, this provides evidence against the strategic account. This finding was as anticipated given that explicit prioritization had no impact on RI with working memory consolidation. The trial-level findings suggests that VWM consolidation speed could vary from trial to trial; when it is fast, it exerts less PI because it is more likely to be complete before T2. This fast consolidation shows up in Figure 13 as having greater slope of the RI function because there is greater WM consolidation achieved before T2 onset. Bigger RI slopes would thus be associated with faster T2 decisions because they reflect faster WM consolidation and thus less PI (expressed by faster T2 RTs).

One limitation of the current study involves the between-subjects comparisons. Data from only 16 participants were collected for each study in order to match the number used in the original Nieuwenstein and Wyble (2014) RI study. A power analysis suggested that this was sufficient for the within-subjects analysis, however, may be underpowered for between-subjects comparisons. Although there is reduced power for the between-subjects effects, Bayes factors do not converge towards a particular result as a function of sample size (this is not the case for null hypothesis significance test; Rouder, Speckman, Sun, Morey, & Iverson, 2009). This approach does not give advantage to rejection of the null, instead being able to determine if there is strong evidence in favor, against, or not enough evidence to know (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017;

Vandekerckhove, Rouder, & Kruschke, 2018; Wagenmakers et al., 2018). It is for this reason why Bayes factors were included for the between-subjects comparisons. Another limitation of the present study is that only one experiment was done using non-verbal stimuli (i.e. color patches). Because of this, it cannot be absolutely certain that all findings observed in the other 8 experiments with visually presented verbal stimuli also apply to non-verbal stimuli.

In conclusion, the present research provides further support that the slower estimate of consolidation rate revealed from the retroactive dual-task interference paradigm is a structural feature of WM encoding, not strategic. The WM consolidation rate may be slower than originally anticipated, challenging decades of research which has relied on the assumption that WM consolidation is rapid, and that visual masking can be used to terminate what is believed to be a unitary consolidation process. The discrepancy of WM consolidation estimates may be due in part to the possibility that the various approaches are measuring separate stages of consolidation, rather than the whole process. This thesis characterizes RI on WM consolidation by a second task. Future work will continue to evaluate what aspects of T2 are sufficient to drive RI. Planned future research outside the scope of this thesis will investigate whether there is a relationship between individual differences in WM storage capacity, and the amount of proactive and/or retroactive interference obtained.

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. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Alvarez, George A., & Cavanagh, P. (2008). Visual short-term memory operates more efficiently on boundary features than on surface features. *Perception & Psychophysics*, *70*(2), 346–364. https://doi.org/10.3758/pp.70.2.346

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). https://doi.org/10.1016/S0079-7421(08)60452-1

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*(10). https://doi.org/10.1167/11.10.6

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38–70. https://doi.org/10.1037/0033-295X.114.1.38

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595. https://doi.org/10.3758/BF03212157

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547.

Capa, R. L., Bouquet, C. A., Dreher, J.-C., & Dufour, A. (2013). Long-lasting effects of performance-contingent unconscious and conscious reward incentives during cued task-switching. *Cortex*, *49*(7), 1943–1954. https://doi.org/10.1016/j.cortex.2012.05.018

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127. https://doi.org/10.1037/0096-1523.21.1.109

Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: resource depletion or temporary loss of control? *Psychological Research*, *69*(3), 191–200. https://doi.org/10.1007/s00426-004-0173-x

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133. https://doi.org/10.3758/BF03206454

Erkal, N., Gangadharan, L., & Koh, B. H. (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review*, *101*, 528–545. https://doi.org/10.1016/j.euroecorev.2017.10.021

Fuller, R. L., Luck, S. J., McMahon, R. P., & Gold, J. M. (2005). Working memory consolidation is abnormally slow in schizophrenia. *Journal of Abnormal Psychology*, *114*(2), 279–290. https://doi.org/10.1037/0021-843X.114.2.279

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 845–866. https://doi.org/10.1037/0096-1523.19.4.845

Greenhouse, I., & Wessel, J. R. (2013). EEG signatures associated with stopping are sensitive to preparation. *Psychophysiology*, *50*(9), 900–908. https://doi.org/10.1111/psyp.12070

Hao, R., Becker, M. W., Ye, C., Liu, Q., & Liu, T. (2018). The bandwidth of VWM consolidation varies with the stimulus feature: Evidence from event-related potentials. *Journal of Experimental Psychology. Human Perception and Performance*, *44*(5), 767–777. https://doi.org/10.1037/xhp0000488

Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Kahneman, D. (1973). *Attention and effort*. https://doi.org/10.2307/1421603

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking-An integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*(6), 557–583. https://doi.org/10.1037/bul0000144

Lagroix, H. E. P., Spalek, T. M., Wyble, B., Jannati, A., & Di Lollo, V. (2012). The root cause of the attentional blink: first-target processing or disruption of input control? *Attention, Perception & Psychophysics*, *74*(8), 1606–1622. https://doi.org/10.3758/s13414-012-0361-5

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(2), 429–438. https://doi.org/10.1037/a0023925

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305. https://doi.org/10.1016/j.tics.2005.04.010

Marti, S., Sigman, M., & Dehaene, S. (2012). A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period. *NeuroImage*, *59*(3), 2883–2898. https://doi.org/10.1016/j.neuroimage.2011.09.063

Miller, J. R., Becker, M. W., & Liu, T. (2014). The bandwidth of consolidation into visual short-term memory (VSTM) depends on the visual feature. *Visual Cognition*, *22*(7), 920–947. https://doi.org/10.1080/13506285.2014.936923

Neisser, U. (1967). *Cognitive psychology*. East Norwalk, CT, US: Appleton-Century-Crofts.

Nieuwenstein, M., & Wyble, B. (2014). Beyond a mask and against the bottleneck: Retroactive dual-task interference during working memory consolidation of a masked visual target. *Journal of Experimental Psychology: General*, *143*(3), 1409–1427. https://doi.org/10.1037/a0035257

Olivers, C. N. L., van der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: against a limited-capacity account of the attentional blink. *Psychological Research*, *71*(2), 126–139. https://doi.org/10.1007/s00426-005-0029-z

Ouimet, C., & Jolicoeur, P. (2007). Beyond Task 1 difficulty: The duration of T1 encoding modulates the attentional blink. *Visual Cognition*, *15*(3), 290–304. https://doi.org/10.1080/13506280600693741

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290. https://doi.org/10.3758/BF03203943

Potter, M. C., Chun, M. M., Banks, B. S., & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: the visual attentional blink and an amodal task-switch deficit. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(4), 979–992.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860. https://doi.org/10.1037/0096-1523.18.3.849

Rieger, T., Mittelstädt, V., Dignath, D., & Kiesel, A. (2019). Investigating limits of task prioritization in dual-tasking: evidence from the prioritized processing and the psychological refractory period paradigms. *Psychological Research*. https://doi.org/10.1007/s00426-019-01250-x

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, 225–237.

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304–321. https://doi.org/10.1037/met0000057

Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2016). Opportunity for verbalization does not improve visual change detection performance: A state-trace analysis. Behavior Research Methods, 49, 853–862.

Schevernels, H., Krebs, R. M., Santens, P., Woldorff, M. G., & Boehler, C. N. (2014). Task preparation processes related to reward prediction precede those related to task-difficulty expectation. *NeuroImage*, *84*, 639–647. https://doi.org/10.1016/j.neuroimage.2013.09.039

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: uncorking the central cognitive bottleneck. *Psychological Science*, *12*(2), 101–108. https://doi.org/10.1111/1467-9280.00318

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *Journal of Experimental Psychology. Human Perception and Performance*, *14*(4), 591–600. https://doi.org/10.1037//0096-1523.14.4.591

Shih, S.-I. (2008). The attention cascade model and attentional blink. *Cognitive Psychology*, *56*(3), 210–236. https://doi.org/10.1016/j.cogpsych.2007.06.001

Song, J.-H., & Jiang, Y. (2006). Visual working memory for simple and complex features: an fMRI study. *NeuroImage*, *30*(3), 963–972. https://doi.org/10.1016/j.neuroimage.2005.10.006

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: a threaded cognition model of the attentional blink. *Cognitive Psychology*, *59*(1), 1–29. https://doi.org/10.1016/j.cogpsych.2008.12.002

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, *14*(1), 1–36. https://doi.org/10.1037/h0073262

Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology. Human Perception and Performance*, *29*(1), 3–18. https://doi.org/10.1037//0096-1523.29.1.3

Tombu, M. N., Asplund, C. L., Dux, P. E., Godwin, D., Martin, J. W., & Marois, R. (2011). A Unified attentional bottleneck in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13426–13431.

Vandekerckhove, J., Rouder, J. N., & Kruschke, J. K. (2018). Editorial: Bayesian methods for advancing psychological science. *Psychonomic Bulletin & Review*, *25*(1), 1–4. https://doi.org/10.3758/s13423-018-1443-8

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, 92–114.

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*(6), 1436–1451. https://doi.org/10.1037/0096-1523.32.6.1436

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance—a review and a theory. *British Journal of Psychology*, *43*, 2–19.

Woodman, G. F., & Vogel, E. K. (2005). Fractionating working memory: consolidation and maintenance are independent processes. *Psychological Science*, *16*(2), 106–113. https://doi.org/10.1111/j.0956-7976.2005.00790.x

Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(3), 787–807. https://doi.org/10.1037/a0013902

Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology. General*, *140*(3), 488–505. https://doi.org/10.1037/a0023612

Ye, C., Hu, Z., Li, H., Ristaniemi, T., Liu, Q., & Liu, T. (2017). A two-phase model of resource allocation in visual working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(10), 1557–1566. https://doi.org/10.1037/xlm0000376

Zedelius, C. M., Veling, H., Bijleveld, E., & Aarts, H. (2012). Promising High Monetary Rewards for Future Task Performance Increases Intermediate Task Performance. *PLOS ONE*, *7*(8), e42547. https://doi.org/10.1371/journal.pone.0042547

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860