ADAMS, RUSSELL LANE. M.A. Remembering When to Better Recall What. (2021)
Directed by Dr. Peter Delaney. 72 pp.

The possible role that temporal context plays in recall has been a recurring theme within the spacing literature. This study aims to determine whether remembering an item's temporal context makes that item more likely to be recalled. Participants were shown lists of spaced and singleton words. The measure of temporal context knowledge was a test on memory-for-when: that is, indicating the serial position of any previous repetition of each word on the list. An accurate guess (plus or minus two positions) was assumed to indicate probable temporal context knowledge. In Experiment 1, a free recall test was used in order to test the effects of knowing temporal context on recall. To examine the impact of having an accurate memory-for-when (MFW), the probability of recall on the final test was assessed as a function of location recall accuracy during the study phase. While items with access to temporal context via MFW were recalled more often than items without MFW, the effect seemed to be primarily driven by the primacy region. Experiment 2 added an incidental learning task to remove this possible confound. The benefits of having MFW persisted in Experiment 2 even without primacy. Together, these findings implicate a role for having access to an item's temporal context in memory where items with strong connection to temporal context are more likely to be later recalled.

*Keywords:* reminding, spacing effect, spacing, temporal context, free recall

REMEMBERING WHEN TO BETTER RECALL WHAT

by

ADAMS, RUSSELL LANE

A Thesis

Submitted to

the Faculty of The Graduate School at

The University of North Carolina at Greensboro

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Greensboro

2021

Approved by

Dr. Peter F. Delaney
Committee Chair

APPROVAL PAGE

This thesis written by Russell Lane Adams has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

| | |
|---|---|
| Committee Chair | Dr. Peter F. Delaney |
| Committee Members | Dr. Michael J Kane |
| | Dr. Christopher N. Wahlheim |

November 10, 2021

Date of Acceptance by Committee

September 2, 2021

Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I: REMEMBERING WHEN TO BETTER RECALL WHAT

## Introduction

One of the fundamental questions about memory is why we benefit from a repetition and why are some repetitions more beneficial than others. For example, repetitions that are spaced farther apart promote retention better than repetitions that appear consecutively, which is known as the *spacing effect* (e.g., Crowder, 1976; Delaney et al., 2010; Dempster, 1989; Hintzman, 1974; Maddox, 2016). Comparatively, a participant is more likely to remember an item after multiple spaced exposures than after only one longer massed exposure (e.g., Carpenter et al., 2012; Hintzman, 1976; Toppino & Gerbier, 2014). In order to benefit from spaced repetitions, though, some theories suggest that people must somehow connect (consciously or unconsciously) the individual repetitions into a representation that supports later recall (Delaney et al., 2010; Hintzman, 2004, 2010; Wahlheim et al., 2014). Moreover, the benefit of a repetition is also dependent on it being recognized as previously shown (Asch, 1969; Bellezza et al., 1975; Melton, 1967; Thios & D'Agostino, 1976; Wahlheim et al. 2014). In each of these cases, some repetitions are more beneficial than others.

In this paper I will investigate whether a third factor that has been suggested several times in the literature plays an important role in determining a repetition's benefit to memory: whether or not access to an item's temporal contextual information matters (Delaney et al., 2010; Howard & Kahana, 2002; Malmberg & Shiffrin, 2005, Sahakyan et al., 2013). Temporal context is defined as information about the time scale in which items are presented (Howard & Kahana, 2002) and has been argued by some authors to be used at retrieval similarly to other forms of context such as environmental context (Sahakyan et al., 2013). Theoretically, temporal context aids recall by serving as a retrieval cue for an item, increasing the likelihood of retrieving it. A person who has information about the temporal context of an item should therefore recall the item more often than a person who has no such knowledge, and similarly, items for which a person has information about temporal context should be better remembered than items for which the same person has no such information. Access to temporal context information when trying to recall repeated items has been proposed in several theories to be crucial to the benefits of a spaced repetition (e.g., Delaney et al., 2010; Howard & Kahana, 2002; Malmberg & Shiffrin, 2005).

The existence of temporal context is assumed in order to make memory theories work but there is no experimental method widely accepted to test whether a given item has temporal context information associated with it or not. The present project draws on earlier literature on memory-for-when to try and change that. Specifically, I will propose that access to an item's temporal context memory can be assessed in part via *memory-for-when,* which will influence later memory for that item. Hintzman and Block (1973) proposed that the specific time is recorded along with the item, which allows for each of the study events to be viewed as separable events. As evidence that people store memory-for-when, they noted that people are able to recall each location of an item shown twice within a list. This suggests that the ability for a person to recall exactly where in a list an item occurred (temporal position) is accomplished by retrieving stored temporal context information. In comparison, temporal order judgments (deciding which of two items came first in a list) could in principle be accomplished through relative familiarity of the items, or by using item-to-item associations, and therefore does not necessarily require the retrieval of temporal context (Unsworth & Brewer, 2009). Access to an item's specific memory-for-when can be used as an index to having that item's specific temporal context. This project attempts to use Hintzman and Block's methods to better understand how context memory is involved in free recall.

## Effects of Sensory and Temporal Context Information on Memory

While studying an item, information beyond the scope of the targeted item is incidentally encoded as *contextual information.* Contextual information may include such things as sensory experiences and internal states such as mood (see Klein et al., 2007; Smith & Vela, 2001). Examples of sensory contextual information would include the smell of the candle burning in the room while you studied or the temperature of the study room. Environmental context, which is tied to a specific location, has also been shown to influence memory (Smith et al., 2018;).

Memory tends to be best when the initial study context and the test context are matched; context acts a retrieval cue for memories which increases recall (for a meta-analysis, see Smith & Vela, 2001). The claim that context plays a role in memory is supported by other, real-world examples (e.g. Hamilton, 2011; Sahakyan, 2010). For example, Hamilton (2011) described anecdotal reports from oral histories that smell could trigger vivid memories (the Proust effect), while Godden and Baddeley (1975) showed that recall was better when the testing and study phase took place in the same environment – in this case, lists learned underwater were best recalled underwater.

In more recent memory literature, contextual information which describes time, or *temporal context*, has been theorized to have an important role in memory retrieval. Studies supporting models which rely on temporal context or contiguity to predict recall (Greene, 1989; Hintzman, 2003, 2016; Howard & Kahana, 2002; Howard et al., 2015) have provided evidence for the benefits of having temporal context. For example, in free recall, people often recall items that were presented near in time to the item they just recalled, a finding known as the temporal contiguity effect (Howard & Kahana, 2002). Moreover, Lohnas et al. (2011) found that in free recall, participants that recalled items in order retrieved more correct items. These findings have been interpreted as evidence that the temporal context of the last item retrieved is an important cue used to generate the next item in free recall. The Search of Associative Memory (SAM) and Retrieving Effectively from Memory (REM) models of memory explicitly separate contextual information from item information (Gillund & Shiffrin, 1984; Shiffrin & Steyvers, 1997; Raaijmakers & Shiffrin, 1981). These theories model how items are retrieved from memory, operating on the assumption that as a memory trace is encoded, the item is encoded along with contextual and associative information. In spaced repetitions, the link between the item and its original context is assumed to be strengthened during repetitions, causing it to be sampled more often than items with weaker connections to context (e.g., Malmberg & Shiffrin, 2005; Sahakyan et al., 2008; Sahakyan & Malmberg, 2018).

Assuming that temporal context plays a role in connecting an item to its surroundings (as suggested by contiguity effects), a good proxy for temporal context would be the ability to recall where an item was presented within a list, or memory-for-when. Contiguity implies that we retrieve temporal context when we think of an item, and then use that as a cue to retrieve subsequent items. Therefore, memory-for-when could be used to directly test the retrieval of context information.

Memory-for-when is a very precise and explicit kind of temporal context recollection. Other implicit types of temporal context information, like having a feeling that an item was closer to the beginning or the end of a list, may also exist. Similarly, less precise explicit information such as that the item was in the first half of the list may also exist. Explicit memory-for-when is sufficient to conclude people have temporal context information, but the absence of memory-for-when does not necessarily preclude that other temporal contextual information is present. Thus, using only memory-for-when as an indicator that temporal context information is available is a conservative test of memory for temporal context information.

## How Temporal Context Explains Empirical Facts About Repetition

Any theory of repetition effects needs to be able to explain why some repetitions are better than others and why unrecognized repetitions lack the same benefit as recognized ones. The purpose of this study is not to provide support for one of these theories over another, but rather to provide support for a general mechanism that is implicated in all of these theories. Temporal context-based theories assume that temporal context has a role in connecting an object to its surroundings and that memory and repetition effects depend on the recognition of a repetition as such. If both of these assumptions are true, then spaced repetitions that have accurate temporal placement / memory-for-when should show greater free recall than items with incorrect placement. There are a variety of different theories of spaced repetition that would all be consistent with this assumption, such as encoding variability (Melton, 1967, 1970), the recursive reminding hypothesis (Hintzman, 2004, 2010), study-phase retrieval (Delaney et al., 2018; Greene, 1989; Hintzman & Block, 1973), and Search of Associative Memory and Retrieving Effectively from Memory theories (Malmberg & Shiffrin, 2005; Murnane & Phelps, 1995; Raaijmakers, 2003). My goal is not to adjudicate between the different theories, but rather to note that memory-for-when being greater for recalled items than unrecalled items is consistent with many popular theoretical accounts of spacing.

### Why Does Spacing Benefit Memory?

In many theories of repetition effects, there is an underlying connection between context information and the ability to recall an item: the more context information that is connected to repetition the more valuable the repetition (Greene, 1989; Hintzman, 2004, 2010; Melton, 1967, 1970). For example, the encoding variability theory claims that each experience is recorded alongside the contextual information and the degree of contextual difference between the study and restudy of an item is what determines the repetition's value, or benefit, (Melton, 1967, 1970). Here the encoded information is believed to serve as connections that can be used to recall said item. Massed items lack the contextual differences found in spaced items and therefore have fewer connections to recall an item (Lohnas et al., 2011; Martin, 1972). Similarly, the reminding account and study phase retrieval theories of context both highlight the retrieval of a previously shown item and the encoding of the change in context between repetitions (Greene, 1989; Hintzman, 2004, 2010). In the reminding account, the act of retrieving an item from memory allows for the encoding of new associations between the item being retrieved and the context of the current setting. This process creates configural representations where the memory of later events also includes the operation of reminding and the earlier event. The embedded act of remembering a previous can facilitate memory for temporal order by linking the recollection of current and past events and creating a distinction between the two (e.g., Wahlheim & Jacoby, 2013; Wahlheim & Zacks., 2019).

**Why is Recognition of Repetitions Important?**

All of the above-mentioned theories depend on additional contextual information being encoded during spaced repetitions. These theories assume that additional connections made during spaced repetitions between an item and different contexts will, in turn, make the repetitions more memorable. While memory-for-when is more specific than simply recognizing an item as old, the ability to recognize an item is a precondition for memory-for-when. Thus, if one fails during repetition to make contact with or recall the earlier presentation, then the benefits of spacing for that repetition will be reduced.

Following this logic, a forgotten repetition or forgotten first appearance will not be able to connect to its unique context and will not provide the benefit of spacing, a claim which aligns with previous findings (Asch, 1969; Bellezza et al., 1975; Melton, 1967; Thios & D’Agostino, 1976). For example, Asch (1969) had participants study and learn word pairs in two separate lists. One-word pair was identified as a “critical pair” and was present in both lists. After learning the second list, participants were asked about recognizing a critical pair from the previous list. Participants who could not identify a critical pair from the first list were classified as a “nonrecognition” subject. There was no difference in recall between the non-recognized critical pair and a pair that was only learned once. Conversely, critical pairs that were recognized as existing in both lists had a superior recall in comparison to items learned once. This finding suggests that if participants cannot recognize a repetition, they will have to relearn it as if it were a new item.

Similar findings were obtained by Thios and D’Agostino (1976), who had participants first read a list of sentences aloud and later asked them to transform a phrase within each sentence from using active language to passive. After both of these tasks were completed, participants had to complete a free recall task. A spacing effect was only obtained in the retrieval conditions where when participants had to actively retrieve an item's first occurrence. Moreover, sentences that were correctly transformed in the restudy phase, which required memory of the first repetition, had better free recall than items that were incorrectly transformed. Participants would not have been able to correctly transform the sentence if they did not have access to the first presentation.

More recently, Hintzman (2001) tasked participants in two experiments with studying word lists. During the memory portion of the test, participants were instructed to indicate an item as being either new or old. If old, participants were then asked to indicate if they remembered specific details about that item or if the item only seemed familiar. In the judgments of frequency task (JOFs), participants were asked to indicate how many times a word was repeated in the previous list.  A similar task was also completed for judgments of recency (JORs) with participants indicating if an item had been previously shown and recording how recent the last presentation of that item was. In both JOFs and JORs, items that were remembered had more accurate memory judgments than their familiar counterparts.

**The Role of Temporal Context**

Based on existing theories of repetition, if temporal context can be assumed to behave like other forms of context, it can be expected that the when people can access to a repetition's temporal context, they should be more likely to recall that item. In any of the above theories, temporal context or memory-for-when can be used to connect a repetition's unique context information to an item.

In terms of encoding variability, the difference in time between repetitions would create additional context which is expected to aid in an item's recall. In the reminding account, access to temporal context may provide a pathway to recall that is easier to access than other associations. The change "in when" from one repetition to the second would account for the additional contextual information (Greene, 1989). The role of temporal context as theorized in study-phase retrieval is also supported by other studies that focused on spaced repetitions versus exposure time and the limits of contextual encoding (Malmberg & Shiffrin, 2005; Murnane & Phelps, 1995). Malmberg & Shiffrin (2005) hypothesized that a limited amount of context can be stored per repetition and that beyond the initial 1-2 s of contact with an item, additional time has little impact on quantity context stored. The value of a spaced repetition over a massed repetition could be explained as spaced learning encoding more retrieval cues. Testing memory-for-when provides a way to gauge the accessibility to an item's temporal context in a way that is more than just recognizing it as old. Accurate memory-for-when would indicate that a person can readily access an item's context and would therefore be more likely to recall that item in comparison to items with limited access to the context.

## The Present Research

In the current experiment, study items were presented twice. When participants were able to successfully identify a repeated item as old, they were then tasked to indicate where that item appeared in the previous list. When the location of an item is identified within two spaces, the participant was considered to have memory-for-when for that item. Theoretically we assume that memory-for-when can be used to index an individual's access to temporal context. As we mentioned earlier, many theories of memory propose that temporal context is one of the cues used to retrieve items in free recall.

In the present study, when items are seen a second time during study, we ask participants to identify memory-for-when (temporal context information) given the item. If they are able to do this, then they should later be able to use the temporal context to help retrieve the items, too. But is using temporal context to retrieve items really the same as using the items to retrieve temporal context? I note that one feature of memory and cued recall is that associations are usually reversible; i.e., if Item A is a cue for Item B, then after study B is usually a good cue for item A (Carpenter et al., 2006; Kahana, 2002; Norton, 2016; Murdock, 1966). If we use temporal context to cue an item and if temporal context behaves like other cues, then we would assume that an item could also be used to retrieve temporal context. I would also assume that if temporal context behaves like any other association, that a person who can accurately place items in a previous list to have greater access the item's temporal context.

The current study aims to test whether having access to memory-for-when increases the benefit of a repetition on a final recall test due to its role as an index of temporal context. Logically a repetition should be more beneficial when its temporal context information can be recalled. The difference in recall between items with memory-for-when and without memory-for-when should be indicative of the benefit of having temporal context for a repetition.

Because contemporary memory models like REM assume that recall is more dependent on context information than is recognition (Shiffrin & Steyvers, 1997), the present study also aimed to compare the effects of having memory-for-when on recall tests. Specifically, I showed participants two lists of words. List 1 contained the first presentation of 25 spaced items and List 2 contained both the second presentation of the spaced items and 25 new singleton items. During List 2, participants were asked to identify whether the item currently shown was present in the previous list or not. If so, then they were asked to identify when the item occurred in the original list (from Position 1 to Position 25). Correct memory-for-when occurred if participants accurately identified an item as old and could accurately recall the location of the item in List 1 to within 2 items of the correct serial position. After List 2, participants took a free recall test. The key hypothesis was that more accurate memory-for-when would lead to higher free recall than items that do not have memory-for-when, while for recognition, there would little difference between those who have and lack memory-for-when.

The existence of temporal context is generally inferred rather than measured directly. Since time is a form of context, I expected that knowing the time that an item appeared on a list would facilitate recall. Consequently, I tested peoples' explicit memory for temporal context information during a repetition and see whether it affects the likelihood of later recall. By testing if a participant can recall information about an item's temporal context information during the study phase, I aimed to predict the likelihood that the participant will be able to later recall that item.

# CHAPTER II: EXPERIMENTS

## Experiment 1

Experiment 1's main goal was to test whether having accurate memory-for-when for repetitions that occur during the study phase would be associated with higher free recall rates compared to items that were recognized but without accurate memory-for-when. Participants were given two lists that consisted with spaced and singleton repetitions and were tasked with identifying if an item had been previously presented and if so, where in the list did the item first appear. After both lists were completed, participants were tasked with recalling as many items from any part of the test as they could. I report how I determined my sample size, all data exclusions, all manipulations, and all measures in the study (Simmons et al. 2010).

### Method

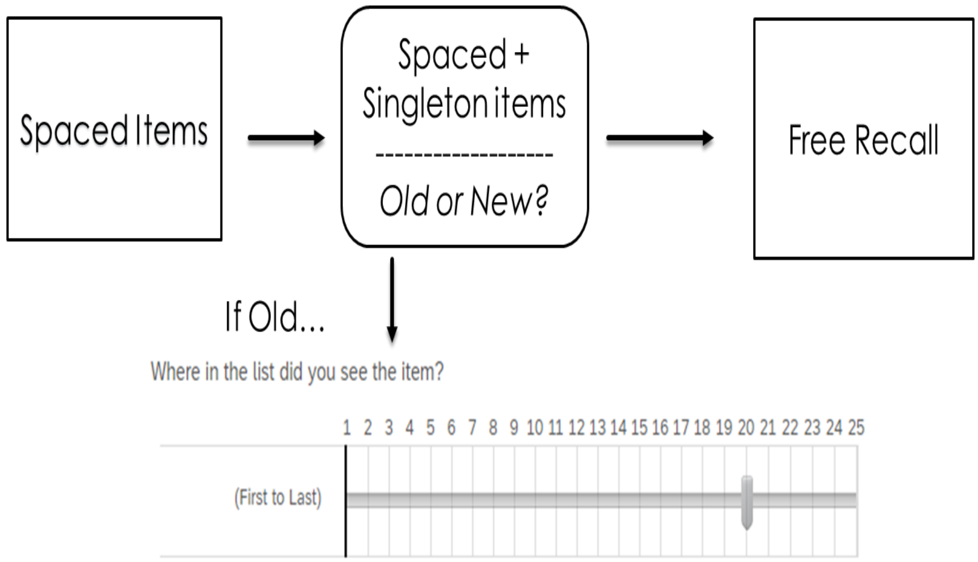
#### *Participants*

Students from the University of North Carolina Greensboro participated in exchange for course credit. Complications due to COVID led to data being collected three different ways: participants were tested individually with 17 participants meeting in person with an experimenter, 8 participants meeting with an experimenter via webcam and 59 participants completing the experiment online without an experimenter. The data from 3 of the online participants were dropped and replaced because they did not complete the task. The data of 84 participants was used in the analysis. Students who had previously participated in a memory-for-when study were not eligible to sign up for this study.

**Materials.** A pool of 50 two-syllable words rated 5 or higher on imagery from the Toronto Noun Pool was divided into two sets of 25 and assigned to serve either as spaced items (which are shown once in the study phase and once in the restudy phase) or singletons items (which were only shown in the restudy phase). The assignment of words to the spaced and singleton conditions, as well as the order in which the words were presented was randomized.

**Procedure.** The study was conducted in three phases: the study phase, restudy phase, and test phase (Figure 1 provides an overview of the procedure used). In the *study phase*, participants were instructed to study the items for a future memory test. The first of two spaced repetitions for 25 items were presented in a random order (the second presentation was later, during the restudy phase). Each word appeared in black print on the center of the white computer screen background for 5 s before automatically transitioning to the next. After studying all items, participants received instructions on how to complete the restudy phase. The instructions used while administering the test can be found in Appendix A.

**Figure 1 Methodology Diagram**




*Note.* Schematic of procedure for Experiment 1.

In the *restudy phase,* the second presentation of all 25 spaced items and the 25 singleton items were intermixed in a random order. Each word appeared in black print on the center of the computer screen for 5 s before automatically transitioning to a set of boxes labeled “new” and “old.” Participants indicated whether they thought the item was shown previously (i.e., “old”) or not (i.e., “new”). When participants checked the “old” box, a slider then appeared and the participant was asked to locate where in the previous list they thought they saw the word (from Position 1 to 25). There was no time limit for participants to check the box or to use the slider but time spent on each question was still recorded. An image of the slider can be found in Appendix B. The slider contained 25 points for each word in the list so that the participants could suggest where in the list they believe the word to have originally occurred. In order to deter knowingly inaccurate guesses, participants were also given the option to check a box labeled “I have no clue." After completing the questions for an item, the next word appeared as before, and the process repeated. Participants were told that this section would be scored, that points would be awarded for correctly identifying an old word, and that extra points would be awarded for correctly identifying where within two items it had appeared during the study phase, with a small penalty for incorrectly guessing the location of the item. The idea of points was introduced to motivate participants, but no points were recorded or analyzed.

In the *testing phase*, participants were given a screen with 50 blank spaces and were instructed to free recall as many words as they could from any part of the experiment. Participants were given 4 m to recall as many words as possible and told that points would be awarded for each correct answer.

***Pilot Study.*** A brief pilot study ($n$ = 7) was run prior to running the main experiment just to verify that there were no floor or ceiling effects. The mean recall was 13.8 and participants did not report confusion with the instructions, so I proceeded with the experiment.

**Power Analysis.** As the effects of memory-for-when at repetition on final recall is unknown, medium/small effects size $d_z$=.4 was chosen as the smallest effect size of interest because I wanted to be conservative using Cohen's (1988) criteria. With *ß* = .95 and $\alpha$ = 0.05, the program G*Power was used to calculate the needed sample size of $n$ = 84 for a within-group *t*-test comparing the final recall proportion for items that had memory for when during the study phase and items not having memory for when during the study phase (Faul et al., 2009). Participants who correctly identified two or fewer items as old during the restudy phase were replaced for noncompliance.

## Results

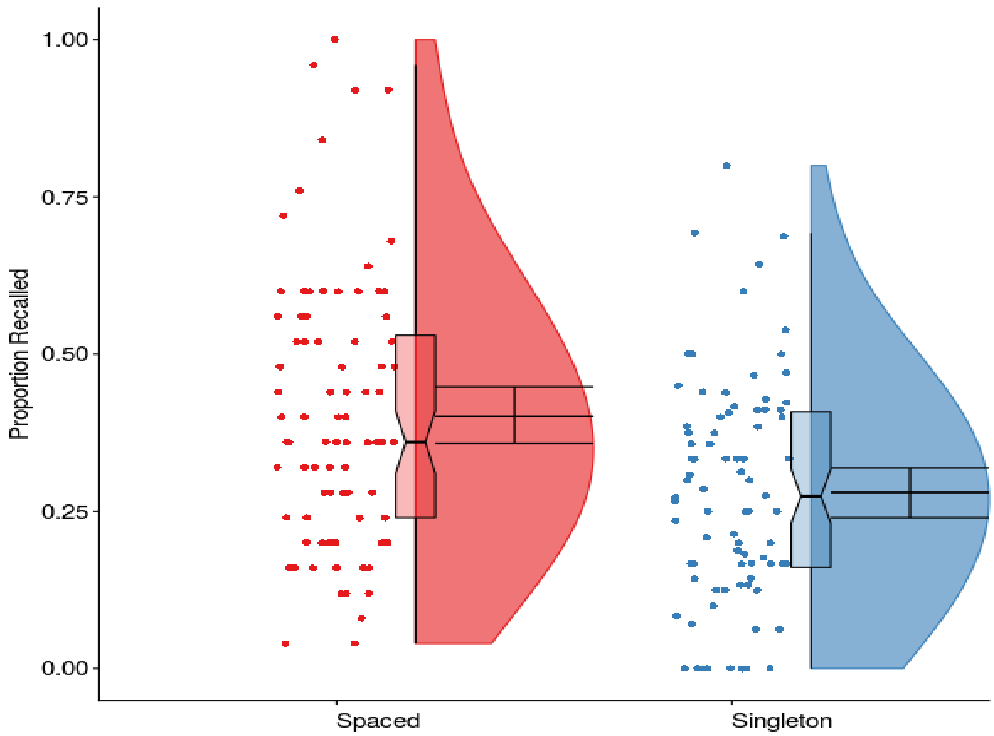
Five planned analyses were run. $\alpha$ was set to .05 for all experiments. Out of all items recalled from study phase, the average proportion of items that had correct memory-for-when was $M = .29$ ($SD = .21$), incorrect memory-for-when $M = .48$ ($SD = .17$), and items incorrectly identified as new $M = .23$ ($SD = .14$).

### *Analyses Related to Final Free Recall*

Free recall was scored as correct if the recalled word matched a word from either the study or restudy phase, allowing for spelling errors and pluralization. Synonyms were not counted as correct.

**Repetition effects on final recall.** Our first analysis addressed the question of whether or not repetition increased recall. A within-subjects *t*-test was used to compare the number of correctly recalled spaced items and singleton items on the final free recall test. The effect of spaced repetition on final free recall is presented in Figure 2. Spaced items ($M = .40$, $SD = .21$) were recalled significantly more often than singletons were ($M = .27$, $SD = .17$), $t(83) = 4.41$, $p < .001$, $d_z = .63$.

**Figure 2 Repetition Effects on Final Recall**



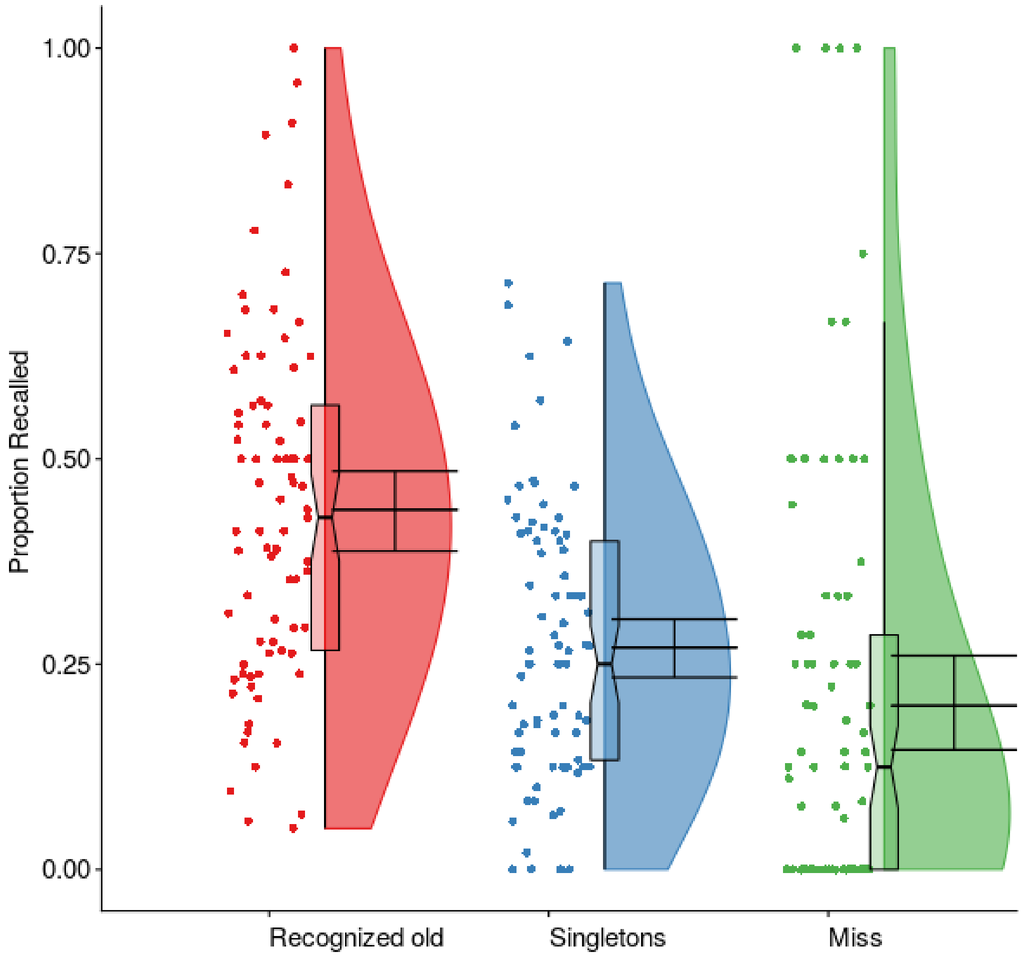

*Note.* Raincloud plot of the proportion recalled on final test as a function of repetition.

**Effects of recognition at repetition during the restudy phase on final recall**.

Past research has shown that the benefit of study repetitions was dependent on recognizing if the items had been previously shown (e.g, Wahlheim et al. 2014). To determine if this was also true in our data, a within-subjects one-way ANOVA was run to compare the proportion of correctly-called-old spaced, incorrectly-called-new spaced (i.e., misses), and singleton words recalled during the testing phase. This represents the benefit of identifying an item as a repetition. If the benefits of repetition on final free recall were dependent on recognizing if an item has been previously shown during the restudy phase, we would expect to see significant differences between the correctly-called-old and incorrectly-called-new conditions. Four participants with a zero-proportion for any of the component item types were dropped to avoid divide by zero errors in the analysis. In the pilot, no participants had a zero-recall rate for any of the item types.

Figure 3 shows the results. The within-subjects one-way ANOVA revealed that the percentage of recalled items differed significantly between correctly-called-old spaced, incorrectly-called-new spaced, and singletons items $F(2,237) = 23.37$, $MSE = .05$ $p < .001$ , $\eta^2 = .10$. Correctly-called-old spaced items ($M = .45$, $SD = .21$) were recalled significantly more often than both incorrectly-called-new spaced items (misses; $M = .19$, $SD = .27$), $t(79) = 8.15$, $p < .001$, $d_z = 1.04$, and singleton items ($M = .27$, $SD = .17$), $t(79) = 5.608$, $p < .001$, $d_z = 0.95$. Singleton items were also recalled significantly more often than incorrectly-called-new spaced items, $t(79) = -2.18$, $p < .05$, $d_z = 0.31$.

**Figure 3 Effects of Recognition at Repetition During the Restudy Phase on Final Recall**

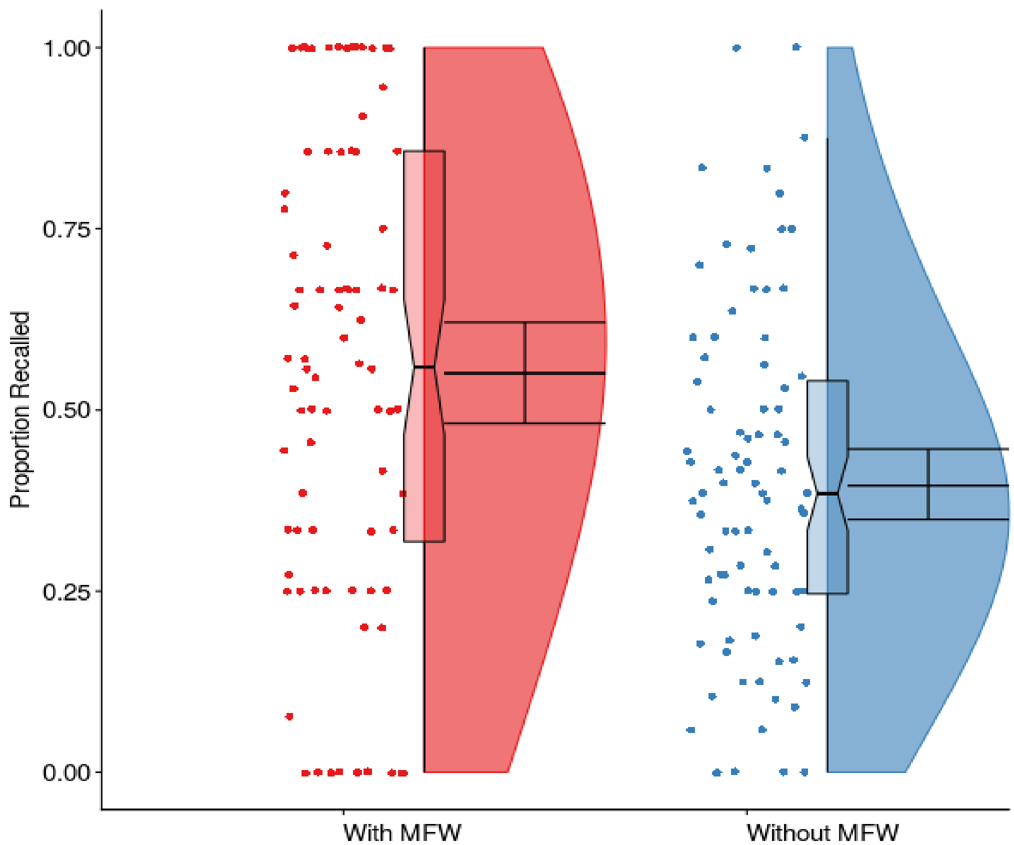



*Note.* Raincloud plot of the proportion recalled on the final free recall test.

**Effects of memory-for-when at repetition on final recall.** The next analysis addressed the critical prediction that a greater proportion of items with memory-for-when would be recalled than items without memory-for-when. A within-subjects *t*-test was used to compare the proportion of correctly recalled items that were recognized as old and either had or lacked memory-for-when (plus or minus two items) and items whose locations in the study phase were not correctly (see Figure 4). Only items that were correctly called old were used in the analysis.

An item was considered to have memory-for-when if the guessed location in the restudy phase falls within two items of the actual location, and incorrect when no guess was recorded or the guess was beyond two items of the actual location. Spaced items with correct memory-for-when ($M$ = .55, $SD$ = .32) were recalled significantly more often than spaced items without correct memory-for-when ($M$ = .39, $SD$ = .24), $t(83) = 3.94$, $p < .001$, $d_z = 0.53$.

**Figure 4 Effects of Memory-For-When at Repetition on Final Recall**



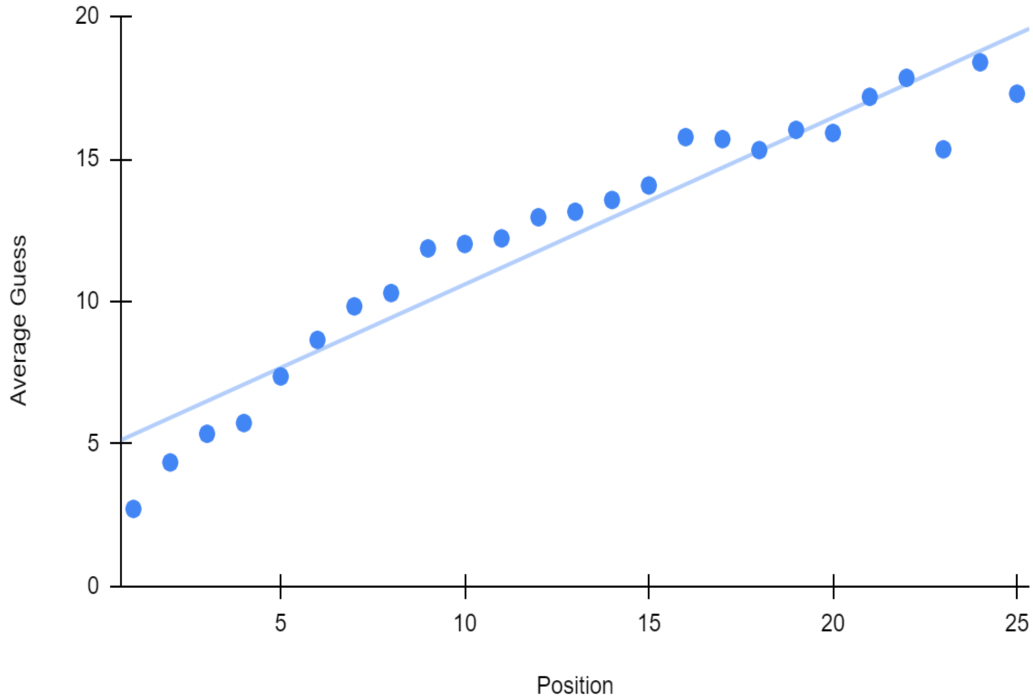

*Note*. Raincloud plot of the proportions of items recognized as old during the restudy phase on final recall.

### *Analyses of Temporal Accuracy*

All previous studies on memory-for-when have followed Hintzman and Block (1971) by reporting the correlation between the input location and the average guessed temporal location for studied items.  Hintzman and Block (1971) also included a plot showing the mean guessed temporal location for each item as a function of serial position at input, including only items that people correctly recalled the temporal location. The mean temporal response was then correlated with the original-list input position. This method weights people who get a lot of hits more heavily, since they provide a larger proportion of the responses. Since this analysis is the original basis for assuming that shows people's memory-for-when judgment are sensitive to the input order, a significant correlation was anticipated. A low correlation would imply that people do not have the ability to recall where an item is in a list. A high correlation would indicate the ability to accurately recall the location of previously shown items.

Figure 5 plots the average guessed position for each item as a function of its serial position. The analysis of temporal location accuracy found a $r(23) = .96$, $p < .001$, correlation between the average position in memory for when judgments for each item and its actual position. However, an autocorrelation analysis of the position to position error found an $r(21) = .62$, $p = .003$, correlation between the error of one position and the next, indicating a significant deviation from linearity.

**Figure 5 Temporal Location Accuracy**




*Note.* Scatter plot of mean position guessed by people indicating memory-for-when as a function of actual serial position.

***Analyses of Output Order***

To test whether some items are recalled earlier on the final recall test, I used a procedure developed by Bjork and Whitten (1974) to calculate output position percentiles. The procedure calculates the average output position during free recall to see if items from one category are output earlier than the items other categories. If items of a particular category are output later than others, we can infer that output interference may impact the recall rates. For each participant, the output percentile was calculated for each item as the average nominal output position of items recalled from each category of item (singletons, repeated/unrecognized, repeated/wrong temporal location judgment, and repeated/correct temporal location judgment) divided by the total number of items recalled by that participant. Thus, higher numbers indicate later output. The data of 21 participants were omitted to avoid zero percentiles for items with memory-for-when and items recognized as old.

The primary question of interest is whether or not items with memory-for-when are output earlier than other items called old during the study phase. The analyses of output order did not indicate that items that with memory-for-when ($M$ = .45, $SD$ = .21) were output earlier than items that solely identified as old ($M$ = .52, $SD$ = .16), $t(62) = 1.67$, $p = .06$, $d_z = 0.38$, although the results were in the predicted direction.

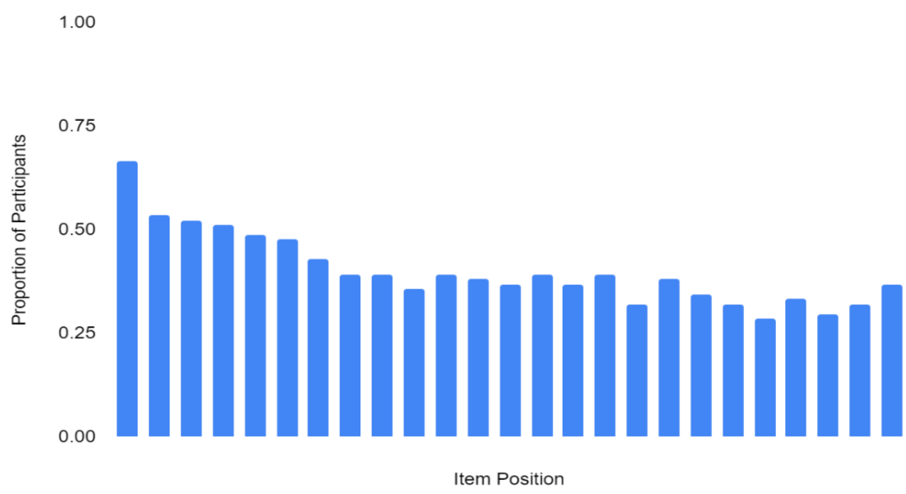
### ***Post-Hoc Analyses***

One concern is that there are primacy effects on memory-for-when accuracy (Delaney et al., 2020; Hintzman & Block, 1971), and there are also primacy effects in memory for words (e.g., Murdock, 1962). Consequently, a possibility worth considering is that any benefits of memory-for-when reflect only the primacy region's general memory advantage.

Serial position curves for all items, as well as items with memory-for-when were plotted to informally examine if the proportion of items with memory-for-when were evenly distributed. Figure 6 shows the serial position function of the study phase. There appeared to be a robust primacy effect during the study phase with little evidence for recency. Figure 7 shows the serial position function for the restudy phase and reflects a recency effect for restudied items. Taken together, these graphs show clear evidence for a primacy effect for the study phase and a recency effect in the restudy phase. The location of the recency and primacy regions across Figures 6 and 7 suggest that people treated the study and restudy phases as one long list, rather than two separate lists, as there was no recency effect in Figure 6 and no primacy effect in Figure 7.

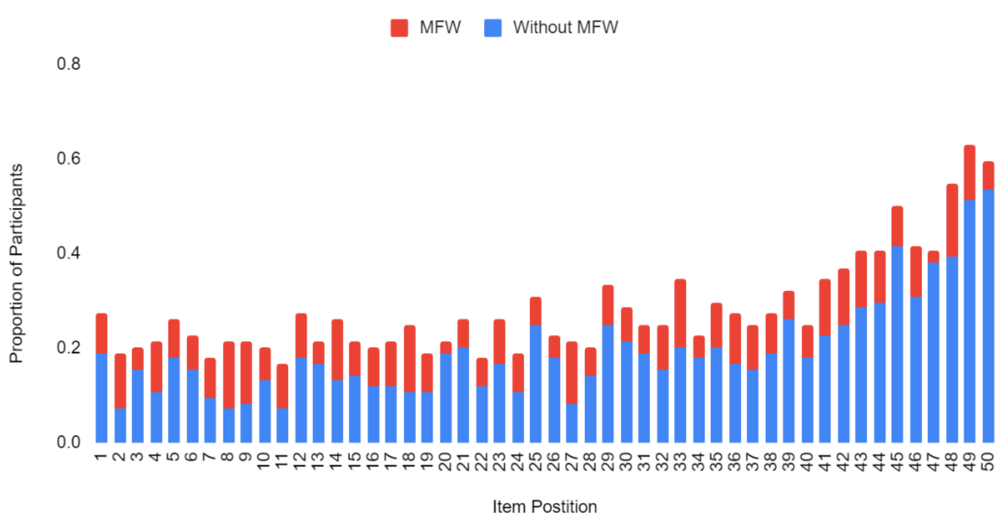
Moreover, items showing memory-for-when were distributed approximately uniformly across Figure 7.

**Figure 6 Serial Position During Study Phase**




*Note.* Bar graph of the proportion of participants who recalled an item at each item position.

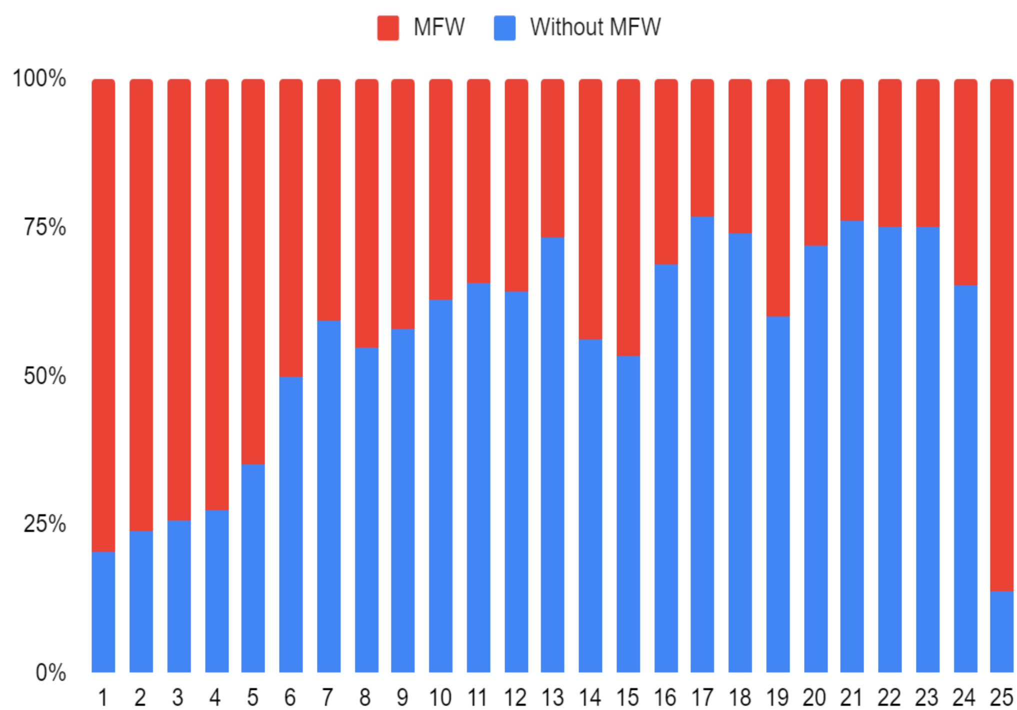
**Figure 7 Serial Position During Restudy Phase**



*Note*. Bar graph of the proportion of participants who recalled an item at each item position during the restudy phase.

Figure 8 shows what percentage of recalled items at each serial position that showed memory for when. The distribution of memory-for-when in Figure 7 shows both a single item recency and robust primacy effect which might suggest that memory-for-when may benefit from factors like recency and primacy.

**Figure 8 Percentage of Recall with Memory-For-When Per Item Position**




*Note.* Bar graph of the percentage of items recalled with memory-for-when at each item position.

Once concern is that any observed benefit of memory-for-when shows primacy and free recall shows better memory for items in the primacy region. Therefore, the benefit of memory-for-when may be due to the confounding of the primacy region which is driving the effect. After removing the first 6 items (which made up the primacy region), a post hoc comparative *t*-test was run between items with memory-for-when ($M$ = .39, $SD$ = .35) and correctly called old items without memory-for-when ($M$ = .39, $SD$ = .27), and no longer found a significant difference in the likelihood of recall, $t(83) = -0.07$, $p = .94$, $d_z = 0.00$.

## Discussion

The primary purpose of Experiment 1 was to show that having memory-for-when (which I argued is associated with temporal context) increased the likelihood of an item being recalled during a free recall test. This result would be expected by theories of memory theories of memory which are context- based and believe that context is used to begin recall. The results of the analysis for memory-for-when at repetition on final recall aligned with our hypothesis that the proportion of items with memory-for-when recalled would greater than items without temporal context. Having access to memory-for-when was linked to higher rates of final recall.

Participants were also found to have good calibration for memory-for-when with a list size of 25. A correlation such as .96 would not have been expected if participants did not have access to memory-for-when and is broadly consistent with theories that suggest that temporal context plays an important role in recall.

Experiment 1 also replicated previous research that if an item is not recognized as a repetition, then it does not receive a benefit from the previous exposure (Asch, 1969; Bellezza et al., 1975; Melton, 1967; Thios & D'Agostino, 1976). However, beyond not receiving a benefit from the first repetition, missed items were found to suffer in comparison to baseline singleton items. This finding contradicts the both early spacing literature (Asch, 1969) and current spacing theory such as the study-phase retrieval account (Hitntzman & Block, 1973) by suggesting a cost for failing to recognize an item. One explanation would be the possibility of a selection bias where items that were easier showed memory-for-when and missed words were more difficult. If memory-for-when was driven by an overall selection bias, then certain words would be expected to be recalled more often than others. Differences in item difficulty may also be idiosyncratic where some items may be easier to remember than others for a given individual. It may be coincidental that the previous literature did not find a cost for missed repetitions if averaging hard and easy items led an insignificant value. If a cost for failed repetition does exist, and it has been previously wiped out by due to list length or multiple repetition, Experiment 1 may have been better able to detect such a cost.

A second explanation for the underperformance of missed items would be interference. Repetition in the study phase can trigger reminding that will increase the probability that an item is later recalled. In A-B, A-D pairs, participants experience interference that lowers recall rates if they are not able to recollect an earlier detected change in response (Wahlheim & Jacoby, 2013). In this study, one could think of the item-context pairs during study and restudy as A-B/A-D pairs, where the cue is the item and the response is the time. Failing to detect the change between the study and restudy opportunities then would create two representations of the item and associated contexts. When one later uses context cues to retrieve the item, these two instances may compete with one another, creating interference.

Post hoc analysis excluding memory-for-when items from the primacy region suggested that the effect of temporal context on recall may be due merely to the primacy effect. Once the first six items of the primacy region were removed, spaced items that had memory-for-when were no longer recalled significantly more often than spaced items without memory-for-when. The extent that memory-for-when was a predictor of recall was dependent on the primacy region which suggests that the benefit of memory-for-when may be a primacy effect in disguise. Specifically, the primacy region is usually attributed to recency and rehearsal (e.g., Rundus, 1971; Tan & Ward, 2000), and those items might show enhanced memory-for-when and recall because they were rehearsed frequently, not because temporal context memory causes recall. An important paper by Tan and Ward (2000) suggested that primacy is caused in part by recency, however, as rehearsal repeats items frequently and at the end of the study period. If so, many have argued that recency depends on temporal context memory (e.g., Howard, Kahana, & Wingfield, 2006), and therefore the result would be consistent with my hypothesis.

Contrastingly, the source monitoring literature has often argued that item memory can be independent from source memory (see Johnson et al., 1993 for a review) and posits that a source judgment is made *after* an item has been identified as old. In memory-for-when, the temporal context aspect of an item's source judgment is used to identify an item as old and to cue its recall. While memory-for-when is a form of source memory, item memory cannot operate independently from its temporal cue. Experiment 2 was designed to remove rehearsal and thereby unconfound rehearsal frequency and temporal context information.

## Experiment 2

Experiment 1 provided preliminary evidence that memory-for-when benefits recall, but results from the post-hoc analysis created reasonable concern that this benefit was driven entirely by primacy. One possible solution would be to introduce an incidental learning task to replace intentional learning from Experiment 1 with incidental learning. Previous studies (Mechanic, 1962; Mechanic & D'Andrea, 1966; Mechanic & Mechanic, 1967) have found that a main difference between the incidental and intentional learning tasks is how the instructions affected the amount of rehearsal of the material. Participants in incidental learning conditions did not expect to be tested and so were less likely to rehearse the material. I expected that removing rehearsal through the use of an incidental learning task would yield a serial position curve with an extended recency effect. The presence of extended recency would align with findings by Marshall & Werder (1972) which sought to observe how the elimination of rehearsal would impact primacy and recency effects. Marshall & Werder (1972) used a word coding task during the presentation of word lists to remove the ability for participants to rehearse. Once rehearsal was eliminated, so too was any benefit of primacy on final recall. Pleasantness ratings have been used before to test the effects of incidental learning (e.g., Geiselman et al. 1983; Sahakyan & Delaney, 2005). If the primacy effect is -- as Tan and Ward (2000) have argued -- largely driven by more frequent rehearsals of items in the primacy region (see also Murdoch, 1962), the addition of an incidental learning task where participants are asked to gauge an item's pleasantness should remove the primacy effect and allow for a view of memory-for-when that is driven solely by recency and not primacy.

## Method

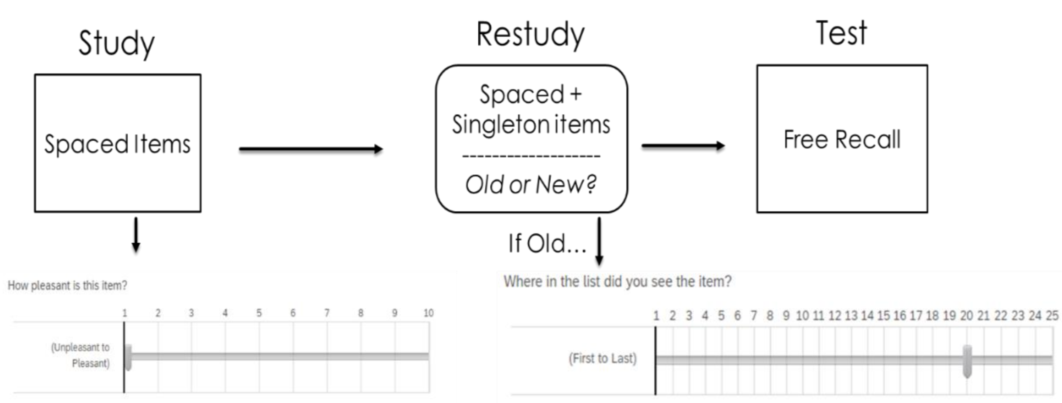
**Participants.** Students from the University of North Carolina Greensboro participated in exchange for course credit. Students who have previously participated in a memory-for-when study will not eligible to sign up for this study.

Although the effect size of the first experiment was found to be $d_z = 0.54$, the benefit of memory-for-when at repetition on final recall without a primacy effect is unknown. Therefore, the same sample size ($n = 84$) was used in Experiment 2. The data from 7 participants who did not complete the restudy task, as defined by correctly identifying two or fewer items as old or did not follow instructions for final recall were replaced.

***Materials***. The same word pool used in Experiment 1 was also used in Experiment 2. Unlike Experiment 1, all participants were tested online (due to COVID-19).

***Procedure.*** The only deviation from the procedure used in Experiment 1 was a change in instructions to convert the study phase into an incidental learning task. When an item was presented during the study phase, participants were tasked with rating the item for pleasantness via 10-point slider (Figure 9 for a graphic representation). After an item had been rated for pleasantness, the next item in the list appeared and the process repeated until the end of the study phase. There was no indication that they would be tested on the words later. The restudy phase was structured in the same as in Experiment 1 with participants using a slider to indicate where in the first list an item indicated as "old" fell.

**Figure 9 Methodology Diagram**




*Note.* Schematic of procedure for Experiment 2.

**Results**

Experiment 2 repeated the same five analyses as Experiment 1. In two of the analyses, repetition effects on final recall and effects of recognition at repetition during the restudy phase on final recall now include the new incidental learning task for singleton items.

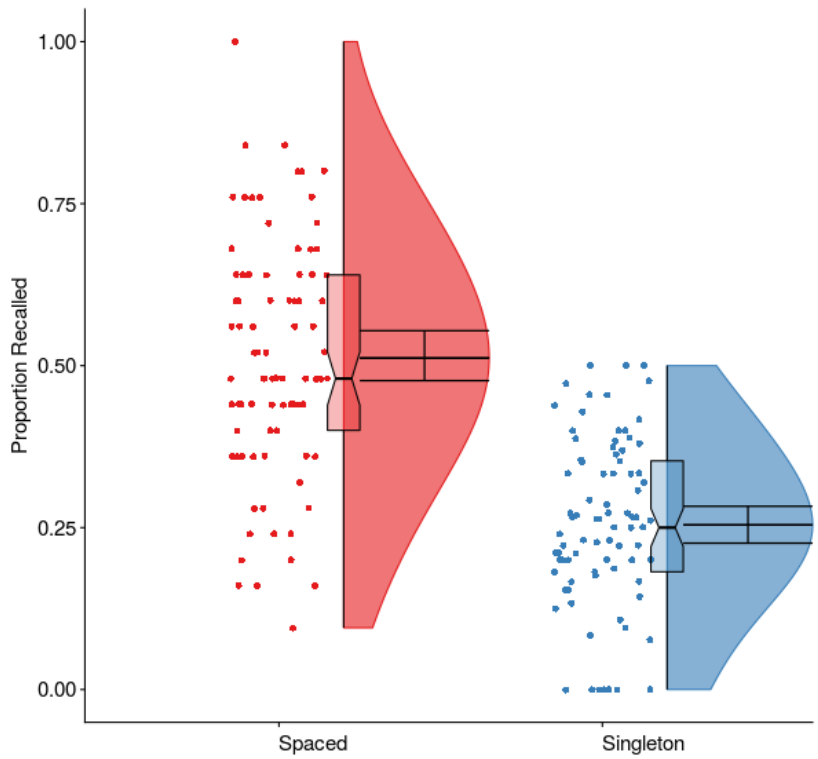
The remaining three, effects of memory-for-when at repetition on final recall, analyses of temporal location accuracy and analysis of output order are the same as in Experiment 1. Alpha was set to .05.

### *Analyses Related to Final Free Recall*

As was with Experiment 1, free recall was scored as correct if the recalled word matches a word from either the study or restudy phase, allowing for spelling errors and pluralization. Synonyms were counted incorrect.

**Repetition effects on final recall.** Our first analysis addressed the question of whether or not repetition increased recall. A within-subjects *t*-test was used to compare the number of correctly recalled spaced items and singleton items on the final free recall test. The effect of repetition on proportion final free recall is presented in Figure 10. Spaced items ($M$ = .52, $SD$ = .18) were recalled significantly more often than singletons were ($M$ = .26, $SD$ = .13), $t(83)$ = 9.96, $p < .001$, $d_z = 1.66$.

**Figure 10 Repetition Effects on Final Recall**

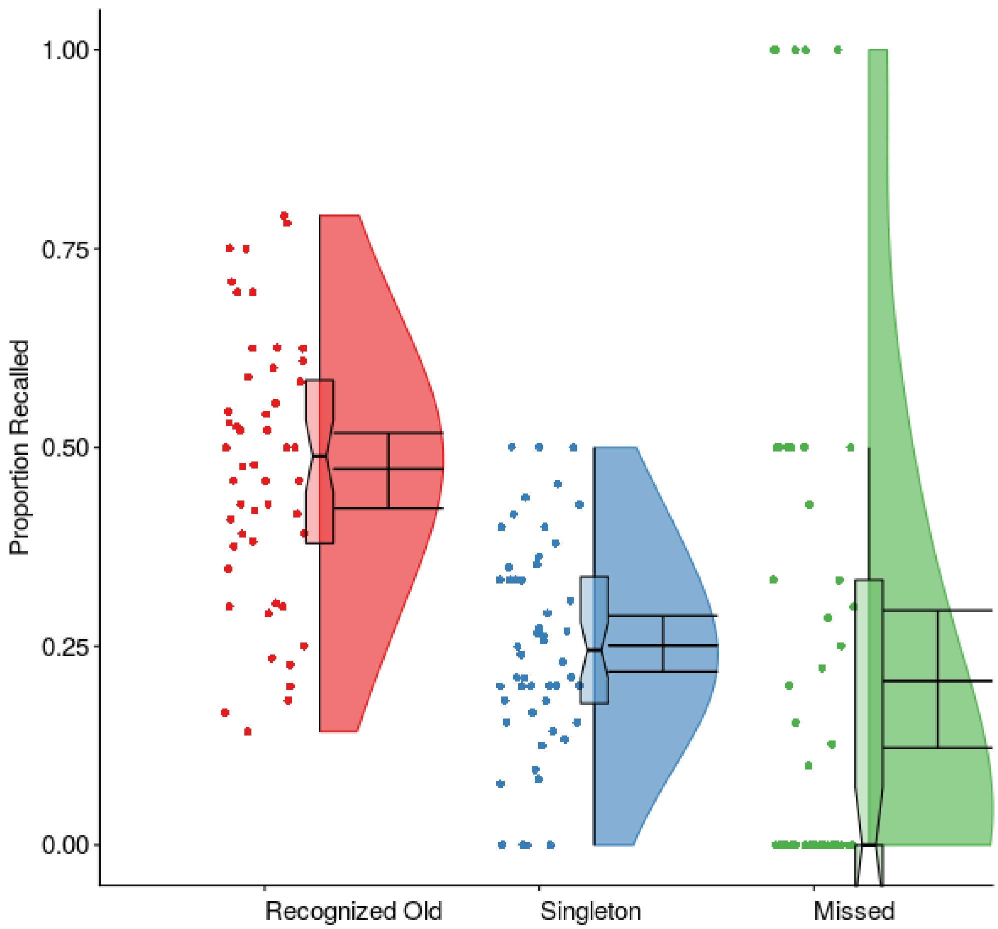



*Note.* Raincloud plot of the proportion recalled on final test as a function of repetition.

**Effects of recognition at repetition during the restudy phase on final recall.** The results from Experiment 1 not only replicated existing findings that items which were not recognized during the restudy would not receive the benefit of the being a repeated item (e.g. Wahlheim et al, 2014), but also that missed items were recalled less often than singletons. The next analysis attempted to detect if this pattern would hold true with the addition of an incidental learning task. A within-subjects one-way ANOVA was conducted to compare the proportion of correctly-called-old spaced, missed (incorrectly-called-new), and singletons recalled during the testing phase. This represents the benefit of identifying an item as a repetition. If the benefits of repetition effects are dependent on recognizing if an item has been previously shown, we would expect to see significant differences between the correctly-called-old and missed conditions. We would also expect that missed items would be recalled at a similar rate to the singletons. If items that were not recognized during the restudy phase performed worse than singleton items, we might suspect interference between the presentations was lowering the likelihood of an item being recalled. Twenty-nine participants with a zero-proportion for any of the component item types were dropped to avoid divide by zero errors in the analysis.

Figure 11 shows the results. The within-subjects one-way ANOVA revealed that the percentage of recalled items differed significantly between correctly-called-old spaced, incorrectly-called-new spaced, and singletons items $F(2,126) = 26.74$, $MSE = .047$, $p < .001$, $\eta^2 = .25$. Correctly-called-old spaced items ($M = .48$, $SD = .17$) were recalled significantly more often than both incorrectly-called-new spaced items (misses; $M = .19$, $SD = .31$), $t(54) = 6.40$, $p < .001$, $d_z = 1.18$, and singleton items ($M = .26$, $SD = .13$), $t(54) = 7.23$, $p < .001$, $d_z = 1.45$. Singleton items were also not recalled significantly different than incorrectly-called-new spaced items, $t(54) = 1.35$, $p = .09$, $d_z = .28$.

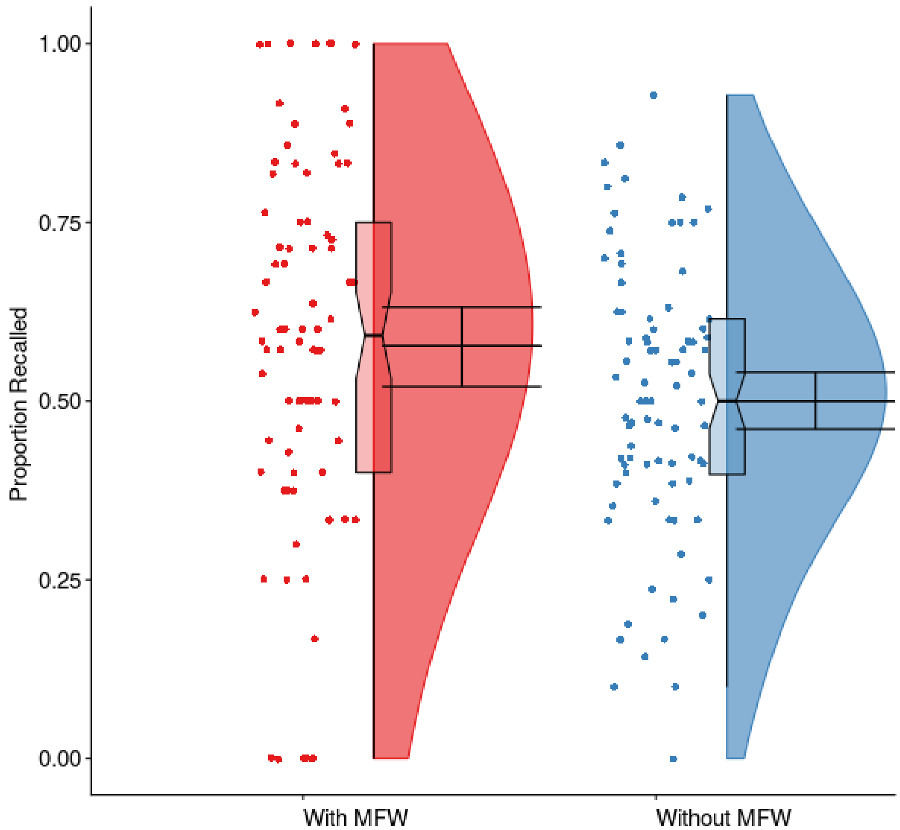
**Figure 11 Effects of Recognition at Repetition During the Restudy Phase on Final Recall**




*Note.* Raincloud plot of the proportion recalled on the final free recall test

**Effects of memory-for-when at repetition on final recall.** The next analysis aimed to replicate the Experiment 1 finding that a greater proportion of items with memory-for-when will be recalled than items without memory-for-when. A within-subjects *t*-test was used to compare the proportion of spaced items correctly recalled on the final free recall test for items whose location in the study phase are correctly identified (plus or minus two items) and items whose locations in the study phase are not correctly identified**.** Only items that were correctly called old were used in the analysis. Spaced items with correct memory-for-when ($M$ = .58, $SD$ = .26) were recalled significantly more often than spaced items without correct memory-for-when ($M$ = .50, $SD$ = .19), $t(83) = 2.66$, $p$ =.009, $d_z$ = 0.34. See Figure 12 for means and distributions.

**Figure 12 Effects of Memory-For-When at Repetition on Final Recall**



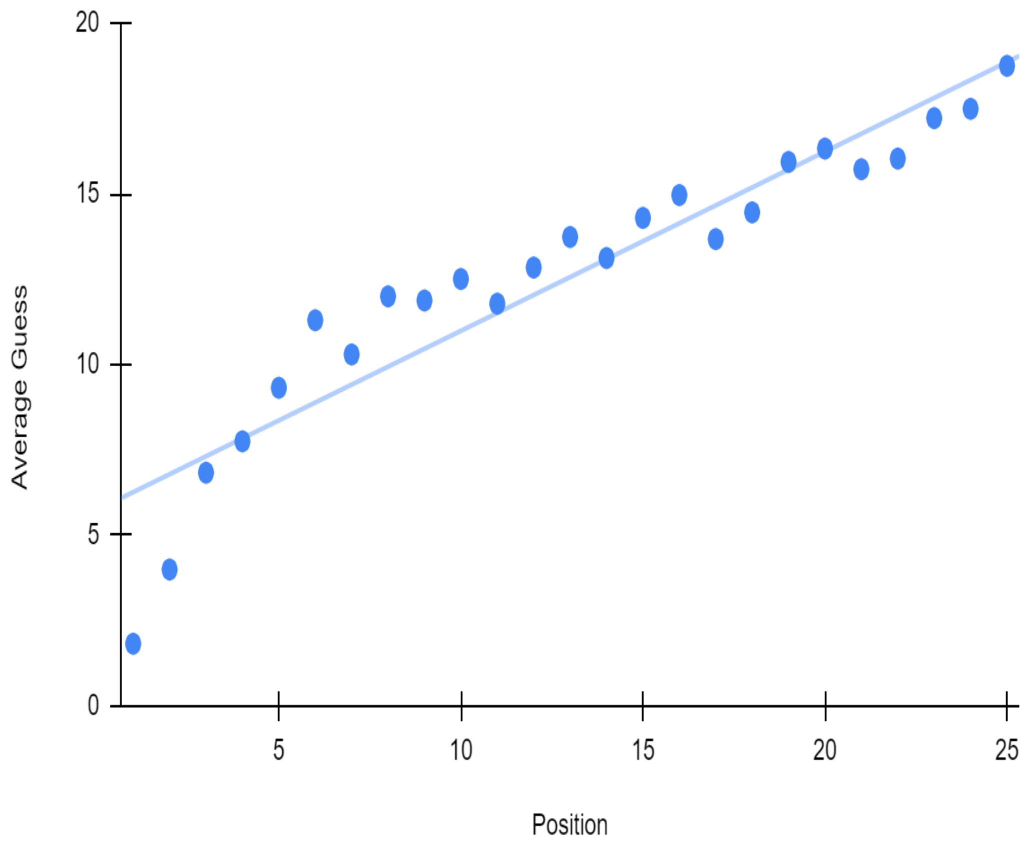

*Note.* Raincloud plot of the proportions of items recognized as old during the restudy phase on final recall.

***Analyses of Temporal Accuracy***

Experiment 2 also aimed to replicate the findings of Experiment 1 in reporting the correlation between the input location and the average guessed temporal location for studied items. Experiment 2 again followed Hintzman and Block's (1971) method of correlating each temporal response with the original-list input position. A low correlation would imply that people lack the ability to recall where an item is in a list. A high correlation would indicate the ability to accurately recall the location of previously shown items.

Figure 13 plots the average guessed position for each item as a function of its serial position. The analysis of temporal location accuracy found a high correlation between the average position in memory for when judgments for each item and its actual position, $r(23) = .93$, $p < .001$,. An autocorrelation analysis of the position to position error found an $r(21) = .29$, $p = .17$, correlation between the error of one position and the next, providing no evidence of deviation from linearity.

**Figure 13 Temporal Location Accuracy**




*Note.* Scatter plot of mean position guessed by people indicating memory-for-when as a function of actual serial position.

***Analyses of Output Order***

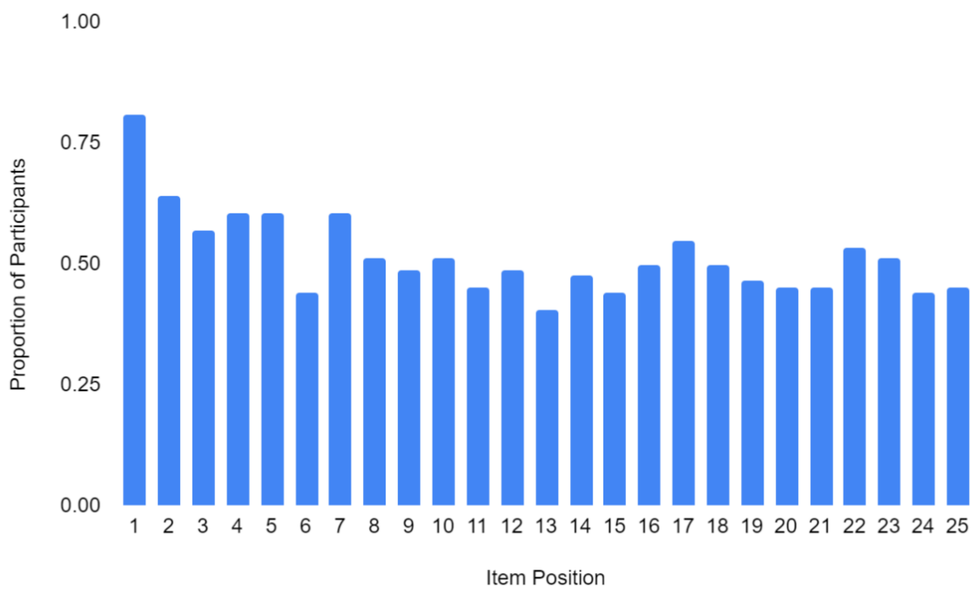
To test whether items with memory-for-when were recalled earlier on the final recall test, I used a procedure developed by Bjork and Whitten (1974) to calculate output position percentiles. We would expect this to be true if memory-for-when represents context strength.

The primary question of interest is whether or not items with memory-for-when are output earlier than other items called old during the study phase. As in Experiment 1, the analyses of output order did not indicate that items that with memory-for-when (*M* = .50, *SD* = .19) were output earlier than items that solely identified as old (*M* = .53, *SD* = .10), $t(75) = 1.31$, $p = .10$, $d_z = 0.24$, although the results were in the predicted direction.

***Analysis of Serial Position***

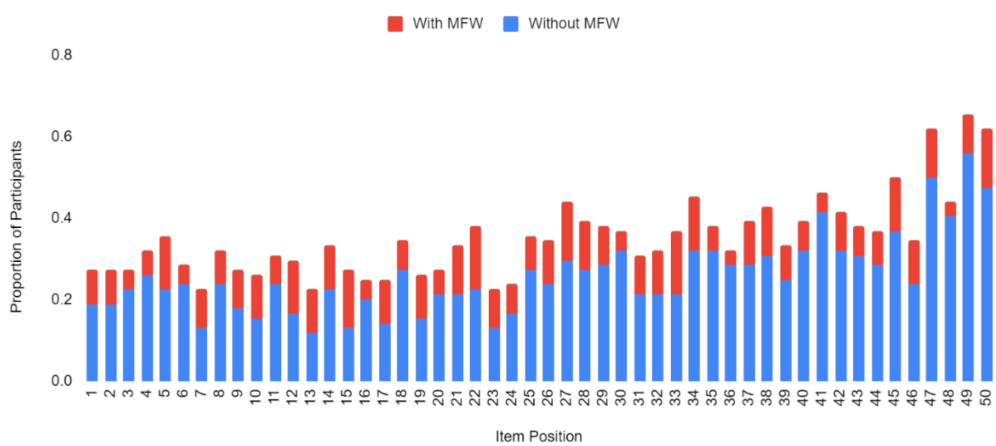
Serial position curves for all items and for the subset of items displaying memory-for-when were plotted to informally examine if the proportion of items with memory-for-when were evenly distributed now that temporal information was encoded incidentally. Figure 14 shows the serial position function of the study phase and gives evidence of at least a single item primacy. Figure 15 shows the serial position function for the restudy phase and reflects a recency effect for restudied items. As was with Experiment 1, items showing memory-for-when were distributed approximately uniformly across Figure 15.

**Figure 14 Serial Position During Study Phase**




*Note.* Bar graph of the proportion of participants who recalled an item at each item position

**Figure 15 Serial Position During Restudy Phase**




*Note.* Bar graph of the proportion of participants who recalled an item at each item position during the restudy phase.

Once concern derived from these Figures is that any observed benefit of memory-for-when shows primacy and free recall shows better memory for items in the primacy region. Therefore, the benefit of memory-for-when may be due to the confounding of the primacy region which is driving the effect. After removing the first 6 items (which made up the primacy region), a post hoc *t*-test was run to compare items with memory-for-when ($M$ = .39, $SD$ = .35) and correctly called old items without memory-for-when ($M$ = .39, $SD$ = .27), which did not produce a significant difference in the likelihood of recall, $t(83) = -0.07$, $p = .94$, $d_z = 0.00$.[1]

[1] A graphic representation of memory-for-when at different range tolerances (perfect, ±2, ±5) for both experiments has been made available in the appendices. Tables of Means and Standard Deviations for both experiments are also available in the appendices.

# CHAPTER III: DISCUSSION AND CONCLUSION

## Discussion

The primary purpose of Experiment 2 was to test whether removing rehearsal would eliminate the benefit of having memory-for-when. This was accomplished by replacing the intentional study task from Experiment 1 with an incidental learning task in Experiment 2. I argued that removing rehearsal of temporal context by introducing a pleasantness rating task would yield a serial position curve with an extended recency effect. The recency region in Experiment 2 (Figure 15 did extend further than in Experiment 1 (Figure 6B) beginning 14 items earlier than in Experiment 1.

A post-hoc analysis excluding the first six items -- which corresponded to the primacy region in Experiment 1 -- suggested that even with the introduction of an incidental learning task, any benefit to having memory-for-when was tied to the primacy region. Once the first six items were removed, spaced items that had memory-for-when were no longer recalled significantly more often than spaced items without memory-for-when. As was with Experiment 1, the extent that memory-for-when was a predictor of recall was dependent on the primacy region.

Experiment 2 did not replicate the finding that unrecognized repetitions were less likely to be recalled in comparison to baseline singleton items. One possibility for the failure to replicate would be that the number of dropped participants (25 more than in Experiment 1) reduced the ability to detect an interaction. The shape of and direction of the data for missed items is similar across both Experiments 1 and 2 even though an interaction was detected for only Experiment 1.

## General Discussion

This work was motivated by an attempt to observe the effects of recollecting an item's temporal context when it is repeated. Temporal context plays a central role in many contemporary theories of memory, such as TCM (Howard & Kahana, 2002), SAM/REM (Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997), and some theories about the benefits of spacing and testing (Lehman et al. 2014; Roediger & Karpicke, 2011) and directed forgetting (for recent reviews, see Delaney et al., 2020; Sahakyan et al., 2013). However, it has not been possible to directly assess whether people have stored mental context or not, because it is typically an unobservable theoretical entity whose existence can only be assumed, not observed. The present paper attempted to change that by asking people to provide memory-for-when judgments for items they had previously studied.

Perhaps the most important difference in the design of the current study from that of previous experiments involving temporal judgments is the use of MFW (see Hintzman; 2001, 2010). Contextual associations are thought to serve as cues that aid in the navigation to and retrieval of items from memory.  Findings from Hintzman (2001) show that we are sensitive to temporal context memory or at least the difference in being able to recall an item versus being familiar.  When an item is presented, it is encoded along with distinct temporal information that can be used as a cue to distinguish between multiple presentations of the same item.

The present study used a measure of MFW and tested the prediction that an item with access to temporal context via MFW should have higher rates of later recall. Experiment 1 tasked participants to report if an item was old or not, and if it was old, then they made a judgment of when in the first list the item appeared. As I predicted, items for which participants had MFW outperformed items that were correctly identified as old but that lacked MFW.

However, one concern was that most of the items with MFW in Experiment 1 fell within the primacy region. This raised the possibility that the relationship between MFW and later recall was spurious and was driven solely by the confound of having both better MFW for early items on the list (thanks to the distinctiveness of those items), and having higher recall for the primacy region (thanks to rehearsal). Consistent with this explanation, once the primacy region was removed the statistical difference between having temporal access via MFW was no longer significant. However, even with the removal of rehearsal, Experiment 2 showed that the benefit of having MFW was limited to the primacy region. Experiment 2 introduced an incidental learning task that eliminated rehearsal, and so should have mostly eliminated the portion of the primacy effect driven by rehearsal (cf. Wixted & McDowell, 1989). Even when encoded incidentally, items that had accurate access to their temporal context received approximately the same benefit to later recall that was seen in Experiment 1.

Surprisingly, the results of Experiment 2 showed that the benefit of MFW was still tied to and did not extend past the primacy region even after removal of rehearsal. The persistence of primacy in the second experiment is consistent with earlier findings by Tan and Ward (2000) which suggest that primacy items have an advantage in recall due to temporal distinctiveness. If, as argued by Tan and Ward, a strong attachment to context leads to better recall, we would expect the primacy region to only consist of items that show a strong connection to context. When rehearsal was removed in Experiment 2, the first item showed the greatest boost from primacy and also had the highest proportion of both MFW and recall.

**Relationship to Memory Theories Grounded in Mental and Temporal Context**

The major finding from the current study may help to answer the question of why some repetitions are more beneficial than others and suggests that access to an item's temporal context may be linked its benefit. The study found that spaced items where a person can recollect the earlier temporal context are more useful than items that are identified as old but do not have access to explicit temporal information. Some theories of spacing suggest that people must be able to create a representation that connects the individual repetitions in a manner that supports recall (Delaney et al., 2010; Hintzman, 2004, 2010; Wahlheim et al., 2014) The benefit of previous repetitions has previously been shown to be dependent on the recognition of the repetition as being old (Asch, 1969; Bellezza et al., 1975; Melton, 1967; Thios & D'Agostino, 1976; Wahlheim et al. 2014). Taken together, Experiments 1 and 2 further suggest that explicit access to an item's temporal context provides a benefit to memory that is above and beyond simply recognizing an item as old.

Context-based theories of repetition suggest that an item is linked to contextual information that is encoded along with the target item. Temporal context specifically has been theorized to have a central role during retrieval, as we often use temporal context as a cue to distinguish relevant and current information from older, less-relevant information. An item's specific context can be used for navigation and provides a useful way to search memory. Hintzman and Block (1971, 1973) suggested that the occurrence of a spaced repetition is represented in memory by a distinctive time tag or retrievable quantity. The existence of unique temporal markers for each item/repetition would explain peoples' ability to mentally distinguish and later recollect multiple presentations of the same item. If an item existed in memory as a single presentation which was constantly updated, we would not expect a person to be able to recall specific aspects of separate repetitions of the item, yet in these studies, people were able to differentiate between repetitions and retrieve context information that was unique to a specific presentation. My finding that items with MFW show better memory than items recognized as old but lacking MFW is consistent with Hintzman and Block's (1971) assertion that people can distinguish repetitions on the basis of the time at which they were encountered.

The implied role of temporal context in free recall would also be consistent with several theoretical accounts of repetition effects such as TCM (Howard & Kahana, 2002), and SAM/REM (Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997). Both of these theories predict that items with a greater connection to temporal context are more likely to be recalled. Unlike other forms of context information such as the color of a word or its background, access to an item's temporal position is harder to gauge. Hintzman and Block (1971) tasked participants to judge an item's temporal position by showing a 55-item word list. During the testing phase, participants had to first identify an item as old before moving on to making a temporal judgment using 1-10 "zones" to indicate when they thought the item appeared. The zones served to measure by the ability to access temporal context by proxy. If an individual could not remember which zone an item was located it would be reasonable to assume they did not have access to that item's temporal context.

The SAM and REM models provide good explanations for the benefits of having MFW. In SAM (e.g., Malmberg & Shiffrin, 2005; Raaijmakers and Shiffrin, 1981), items are represented as memory traces that contain item, associative, and contextual information. Retrieval of an item is assumed to be based on retrieval cues, including temporal context information. Another assumption made is that the recall of an item is dependent on the cue's associative strength, and when a studied item is repeated people may retrieve, thus strengthening, the trace of a previous presentation. The REM model (see Shiffrin & Steyvers, 1997) separates itself SAM by including that item features and context features are assumed to be paired which would make it possible to retrieve temporal context given the item. One would use the test items features to find a matching item-context pair in memory which would also generate the item's context. The benefit of having MFW can be viewed as the benefit of having a memory trace that contains a strong association with temporal context. Following the assumption that item-context pairs are reversible, access explicit temporal context would also imply access to the associated item.

The ability to measure the quality of an item's connection to its temporal context and the finding that items with MFW are more readily recalled than items without MFW shares some similarities with the *Memory for Change* framework (Jacoby et al., 2015; Wahlheim & Jacoby, 2013; Wahlheim et al., 2019). Wahlheim and Jacoby (2013) posited that when reminded of an item's previous occurrence the information from the first representation is "included" in the second representation.  An item is linked to the context of each repetition. When an item is shown and then subsequently shown a second time the individual is reminded of the first repetition to bridge the two.  In A-B, A-D word pairs, items were more likely to be recalled if the participant was able to detect and recall the change from B to D. Conversely participants experienced interference when unable to detect said change. The findings from the present experiment also observe the trend of facilitation or proactive interference depending on the recognition of a repetition moreover, singleton items were recalled more often than missed spaced items. Wahlheim et al. (2019) found a similar effect when testing the benefits of reminders and memory for change. When participants were unable to detect a change from A-B to A-D, they experienced interference and were less likely to recall the original word pair. What is unique about the present study is rather than having A-B, A-D word pairs where two items are paired together, here an item is competing against itself and the time the item was presented. When participants were able to detect and recollect the temporal change via MFW, the items were recalled more often than items without an accurate temporal recollection which was recalled more than items that did not detect a temporal change.

The idea that repetitions are connected through reminders of previous repetitions is also similar to theories of repetition effects advanced by Delaney et al. (2010) and theories of testing effects advanced by Karpicke et al. (2014). Delaney et al. (2010) proposed that the benefits of spaced repetitions are tied to the retrieval of previous presentations which leads to the encoding of additional context. Therefore, presentations with strong access to context (and diverse context such as long lags in spacing) receive a greater benefit from spacing. Lehman et al. (2014) found support for the episodic context account which assumes that retrieval of a repetition improves the ability to guide the memory search of future tests by altering the representation of episodic context.  Lehman et al. tasked participants with studying multiple word lists and compared the final recall of those with elaborative and retrieval practice conditions and found significantly greater recall for participants who practiced by recalling the words they had just practiced.

Testing MFW is similar to Tuliving's RK paradigm but has advantages over traditional distinctions. In MFW, participants have the ability to prove if they know an item and to make sure that their decision is not just based on familiarity. Tasking participants to indicate where in a list an item appeared before a final recall test gives an explicit non-subjective criterion that would be harder to explain as certain items surpassing the signal strength criteria needed to be marked as R.

Tulving's (1985) remember- know (RK) paradigm was introduced to measure different states of awareness that underlie memory retrieval. Participants were asked during a memory experiment to indicate how they judged a previously studied item. Participants were asked if they “knew” an item was old or if they were able to “remember” details about its earlier presentation. In the *dual-process interpretation* of the RK paradigm, R and K are assumed to reflect different forms of memory retrieval. Jacoby, Yonelinas, and Jennings (1997) proposed that the underlying memory processes in the RK paradigm could be identified as recollection and familiarity. Other researchers have suggested that R and K responses reflect varying levels of confidence within memory retrieval (e.g., Donaldson, 1996; Hirshman, 1998; Inoue & Bellezza, 1998). In the signal detection interpretation, the judgment of “remembering” or “knowing” a previous item is dependent on two sets of criteria. If an item meets the more stringent criterion, then participants give a R response, while items that only pass the less stringent criterion are marked as K.

**Is MFW Just the Testing Effect?**

Whiffen and Karpicke (2017) conducted a study similar to ours aimed at explaining why the testing effect occurs, arguing that it emerges from retrieval of context information during testing. Their study tasked participants to study word lists, and later make forced-choice judgments about which list the item was originally presented in. Participants asked to do this list-discrimination task showed higher recall for tested items compared to items that were merely restudied. List discrimination also had comparable memory benefits to performing other deep encoding tasks during restudy, such as rating pleasantness and making a category judgment. However, items in the list discrimination condition saw the greatest temporal clustering on a final free recall test, suggesting that thinking about which list an item occurred in also influenced the recall of nearby items.

Our results are quite consistent with the episodic context account of testing as well. When we asked people to retrieve not just list membership but specific details about MFW, we found that the items for which they had MFW were recalled far better than those that did not (but were still recognized as old). This extends the findings from Whiffen and Karpicke (2017) by showing that it is not just a between-subjects effect driven by testing (or lack thereof) that occurs on a list-by-list basis; successful retrieval of MFW context information is associated with later recall of the items for which MFW was correctly identified. Moreover, the effect could not be attributed solely to the primacy effect (which was not ruled out in Whiffen and Karpicke [2017]).

**Does Temporal Context Usually Matter for Repetition, or Is It Just Because We Asked?**

While we have assumed that MFW is encoded and used automatically, one obvious concern about the current studies is that we asked participants for explicit judgments of MFW, and perhaps they used this information only because I asked them about it. In the source monitoring framework, recollective aspects of memory can be strengthened during an initial testing phase (Johnson et al., 1993). Brewer et al. (2010) showed that testing specific contextual aspects, such as the gender of a speaker or temporal judgments can enhance recall. However, in the current experiment, the act of testing MFW could not have created the benefit in recall as whatever information was used to generate MFW judgments was already present in the task, and not a new cue that was added later. All items were tested for MFW but only the items with accurate temporal placement received a benefit. If the observed benefits of having MFW can be attributed to simply adding an additional temporal retrieval cue during the restudy phase we would have expected to see a difference in recall.

While my results do not give a definitive answer, there is evidence to suggest that the use of temporal context is obligatory. First, it would be remarkable if participants were not storing and later accessing temporal context considering that they were able to complete the task. They had no warning that they "should" store temporal context, and yet were able to retrieve it. I did not add anything during the initial study phase that would signal a need to encode temporal context, implying that they did so automatically (see also Hintzman & Block, 1971). Secondly, any alternative explanation would need to include the presence/persistence of the primacy region. As mentioned before, one prediction for Experiment 2 was the near-elimination of the primacy region with the introduction of the incidental learning task. Figures 12 and 13A show that the primacy region is still present for at least the first items and then disappears through the list as items are not rehearsed to the degree necessary for a contextual advantage. The persistence of the primacy region without rehearsal aligns with research which has also found evidence that primacy persists to a weak degree even under incidental learning conditions (see Wixted & McDowell, 1989). The change in spacing also agrees with what would be expected based on Tan and Ward's (2000) recency-based account of primacy. According to Tan & Ward (2000), the primacy effect is due to recent and frequent rehearsals of earlier items which give them a contextual advantage. Once rehearsal was removed with the introduction of the incidental learning task, the size of the primacy region shortened which suggests that the effect of primacy may be magnified by rehearsal. The presence of primacy in the current study suggests that that context-based information is still being encoded and also mirrors the early findings by Hintzman and Block (1971) where temporal judgments are the most accurate near the start of the list. Temporal primacy was later argued by Hintzman (2010) as being caused by rapid changes in mental context at the start of the test. Alternatively, participants could be using the start and end of the word list as landmarks to gauge an item's position.

Unlike Whiffen and Karpicke (2017), we did not introduce multiple lists as a cue prior to asking about MFW. This implies that MFW had to be encoded automatically during list study and did not reflect a new cue (list membership) that was made salient during initial study. If MFW were not normally part of our encoding processes, why would it be encoded without knowing that MFW would be tested?

Along the same lines, one final piece of evidence that people use temporal context automatically is the consistency in the benefit of MFW. Experiment 2 used pleasantness ratings during the study phase to make any encoded temporal context incidental. Even with the removal of the primacy region, the proportion of items with MFW recalled remained stable (increasing from $M$ = .55 to $M$ = .58) while the overall proportion of recall for spaced repetitions increased from $M$ = .40 to $M$ = .52 in Experiment 2. This change was driven by an increase in the number of hits without MFW; likely due to the addition of the ratings of pleasantness task which would have provided participants in Experiment 2 with an additional cue to recall the items. If asking people to recall MFW was simply giving a new retrieval cue we would have expected to see a change in the proportion of items with MFW recalled in the incidental learning condition. Instead, the lack of change for those items with an additional cue is consistent with people using it in both studies. A future experiment could further test the effects of a MFW task on recall by directly comparing the results of two groups, one where participants are tasked if identifying an item in the second list as old or new and the other with the old/new distinction plus MFW. This experiment would be able to quantify the benefit, if any, that asking if an item has MFW on recall.

# CHAPTER IV: CONCLUSION

The question of why we benefit from a repetition and why all repetitions are not made equal has been a fundamental question about memory. The overarching goal of the current study was to investigate the role of temporal context and if the connection to an item's temporal context matters. My findings suggest that access to temporal context and the ability to recall "when" an item appeared is associated with an increased likelihood of future recall in the primacy region. The memory advantage tied to MFW captures mainly the primacy region and not recency; which would be expected if the benefit of MFW is simple context effect. Instead, my findings suggest that something is special about the start of the list that shows a memory advantage tied to MFW.

REFERENCES

Asch, S. E. (1969). A reformulation of the problem of associations. *American Psychologist*, *24*(2), 92–102. https://doi.org/10.1037/h0027121

Bellezza, F. S., Winkler, H. B., & Andrasik, F. (1975). Encoding processes and the spacing effect. *Memory & Cognition*, *3*(4), 451–457. https://doi.org/10.3758/BF03212940

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Bireta, T. J., Fry, S. E., Jalbert, A., Neath, I., Surprenant, A. M., Tehan, G., & Tolan, G. A. (2010). Backward recall and benchmark effects of working memory. *Memory & Cognition*, *38*(3), 279–291. https://doi.org/10.3758/MC.38.3.279

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*(2), 173–189. https://doi.org/10.1016/0010-0285(74)90009-7

Block, R. A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(6), 530–544. https://doi.org/10.1037/0278-7393.8.6.530

Brewer, G. A., Marsh, R. L., Meeks, J. T., Clark-Foos, A., & Hicks, J. L. (2010). The effects of free recall testing on subsequent source memory. *Memory*, *18*(4), 385-393.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24*(3), 369-378.

Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review*, *13*(5), 826–830. https://doi.org/10.3758/BF03194004

Chang, Y, Delaney, P.F, & Verkoeijen, P.P.J.L. (2019). The testing effect in immediate recognition. *Journal of Cognitive Psychology*, *31*(8), 825–838. https://doi:10.1080/20445911.2019.1677672

Cohen, J. (1988). Set Correlation and Contingency Tables. *Applied Psychological Measurement*, *12*(4), 425–434. https://doi.org/10.1177/014662168801200410

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing effects in the free recall of unmixed lists. *Journal of Memory and Language*, *52*(1), 120–130. https://doi.org/10.1016/j.jml.2004.09.002

Delaney, P. F., de Leon Guerrero, A. J., Cook, M., & Jones, T. C. (2020). *Remembering when: variation in temporal position judgments for faces, places, and words.* Unpublished Manuscript.

Delaney, P. F., Godbole, N. R., Holden, L. R., & Chang, Y. (2018). Working memory capacity and the spacing effect in cued recall. *Memory*, *26*(6), 784–797. https://doi.org/10.1080/09658211.2017.1408841

Delaney, P. F., Nghiem, K. N., & Waldum, E. R. (2009). The selective directed forgetting effect: Can people forget only part of a text? *The Quarterly Journal of Experimental Psychology*. https://doi.org/10.1080/17470210902770049

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: a deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (pp. 63–147). https://doi.org/10.1016/S0079-7421(10)53003-2

Dempster, F. N. (1989b). Spacing effects and their implications for theory and practice. *Educational Psychology Review*, *1*(4), 309–330. https://doi.org/10.1007/BF01320097

Donaldson, W. (1996). The role of decision processes in Remembering and Knowing. *Memory & Cognition*, *24*, 523–533. https://doi.org/10.3758/BF03200940

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. Journal of Experimental Psychology: General, 112(1), 58.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. https://doi.org/10.1037/0033-295X.91.1.1

Godden, D. R., & Baddeley, A. D. (1975). Context-Dependent Memory in Two Natural Environments: On Land and Underwater. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 371–377. https://doi.org/10.1037/0278-7393.15.3.371

Hamilton, P. (2012). The proust effect: oral history and the senses. *The Oxford Handbook of Oral History*, 218–232.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In *Theories in cognitive psychology: The Loyola Symposium* (pp. xi, 386–xi, 386). Lawrence Erlbaum.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 10, pp. 47–91). https://doi.org/10.1016/S0079-7421(08)60464-8

Hintzman, D. L. (2001). Judgments of frequency and recency: How they relate to reports of subjective awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1347–1358. https://doi.org/10.1037/0278-7393.27.6.1347

Hintzman, D. L. (2003). Robert Hooke's model of memory. *Psychonomic Bulletin & Review*, *10*(1), 3–14. https://doi.org/10.3758/BF03196465

Hintzman, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, *32*(2), 336–350. https://doi.org/10.3758/BF03196863

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, *38*(1), 102–115. https://doi.org/10.3758/MC.38.1.102

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, *88*(3), 297–306. https://doi.org/10.1037/h0030907

Hintzman, D. L., & Block, R. A. (1973). Memory for the spacing of repetitions. *Journal of Experimental Psychology*, *99*(1), 70–74. https://doi.org/10.1037/h0034761

Hintzman, D. L., Block, R. A., & Summers, J. J. (1973). Contextual associations and memory for serial position. *Journal of Experimental Psychology*, *97*(2), 220–229. https://doi.org/10.1037/h0033884

Hirshman, E., & Henzler, A. (1998). The Role of Decision Processes in Conscious Recollection. *Psychological Science - PSYCHOL SCI*, *9*, 61–65. https://doi.org/10.1111/1467-9280.00011

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Howard, M. W., Kahana, M. J., & Wingfield, A. (2006). Aging and contextual binding: Modeling recency and lag recency effects with the temporal context model. *Psychonomic bulletin & review*, *13*(3), 439-445.

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*(1), 24.

Inoue, C., & Bellezza, F. (1998). The detection model of recognition using Know and Remember judgments. *Memory & Cognition*, *26*, 299–308. https://doi.org/10.3758/BF03201141

Jacoby, L. L., & Wahlheim, C. N. (2013). On the importance of looking back: The role of recursive remindings in recency judgments and cued recall. *Memory & Cognition*, *41*(5), 625-637.

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 41*(5), 1282-1297. https://doi.org/10.1037/xlm0000123

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental*

*Psychology: Learning, Memory, and Cognition*, *41*(5), 1282–1297. https://doi.org/10.1037/xlm0000123

Jacoby, L., Yonelinas, A., & Jennings, J. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. *Scientific Approaches to Consciousness*.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, *30*(6), 823–840. https://doi.org/10.3758/BF03195769

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Chapter Seven - Retrieval-Based Learning: An Episodic Context Account. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 61, pp. 237–284). Academic Press. https://doi.org/10.1016/B978-0-12-800283-4.00007-1

Klein, K. A., Shiffrin, R. M., & Criss, A. H. (2007). *Putting Context in Context.* In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (p. 171–189). Psychology Press.

Lehman, M., Smith, M. A., & Karpicke, J. D. (2014). Toward an episodic context account of retrieval-based learning: Dissociating retrieval practice and elaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1787–1794. https://doi.org/10.1037/xlm0000012

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2011). Contextual variability in free recall. *Journal of Memory and Language*, *64*(3), 249–255. https://doi.org/10.1016/j.jml.2010.11.003

Maddox, G. B. (2016). Understanding the underlying mechanism of the spacing effect in verbal learning: A case for encoding variability and study-phase retrieval. *Journal of Cognitive Psychology*, *28*(6), 684–706. https://doi.org/10.1080/20445911.2016.1181637

Malmberg, K. J., & Shiffrin, R. M. (2005). The "one-shot" hypothesis for context storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 322–336. https://doi.org/10.1037/0278-7393.31.2.322

Marshall, P. H., & Werder, P. R. (1972). The effects of the elimination of rehearsal on primacy and recency. *Journal of Verbal Learning and Verbal Behavior*, *11*(5), 649–653. https://doi.org/10.1016/S0022-5371(72)80049-5

Martin, E. (1972). Stimulus encoding in learning and transfer. *Coding processes in human memory*, 59-84.

Mechanic, A. (1962). Effects of orienting task, practice, and incentive on simultaneous incidental and intentional learning. *Journal of Experimental Psychology*, *64*(4), 393–399. https://doi.org/10.1037/h0043060

Mechanic, A., & D'Andrea, J. (1966). Visual and pronouncing responses, and the relation between orienting task and presentations in incidental learning. *Journal of Experimental Psychology*, *71*(3), 343–349. https://doi.org/10.1037/h0022975

Mechanic, A., & Mechanic, J. D. A. (1967). Response activities and the mechanism of selectivity in incidental learning. *Journal of Verbal Learning and Verbal Behavior*, *6*(3), 389-397.

Melton, A. W. (1967). Repetition and retrieval from memory. *Science*, *158*(3800), 532–532. https://doi.org/10.1126/science.158.3800.532-b

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 596–606. https://doi.org/10.1016/S0022-5371(70)80107-4

Melton, A. W., & Martin, E. (1972). *Coding processes in human memory*. Oxford, England: V. H. Winston & Sons.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482–488.

Murdock, B. B. (1966). Forward and backward associations in paired associates. *Journal of Experimental Psychology*, *71*(5), 732–737. https://doi.org/10.1037/h0023113

Murnane, K., & Phelps, M. P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 158–172. https://doi.org/10.1037/0278-7393.21.1.158

Norton, A. (2016). Reversibility in mathematics. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, *27*(3), 431–452. https://doi.org/10.1207/s15516709cog2703_5

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93–134. https://doi.org/10.1037/0033-295X.88.2.93

Roediger III, H. L., & Karpicke, J. D. (2011). Intricacies of spaced retrieval: A resolution. In *Successful Remembering and Successful Forgetting* (pp. 41-66). Psychology Press.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*(1), 63.

Sahakyan, L. (2010). Environmental context change affects memory for performed actions. *Quarterly Journal of Experimental Psychology*, *63*(3), 425–433. https://doi.org/10.1080/17470210903414365

Sahakyan, L., & Malmberg, K. J. (2018). Divided attention during encoding causes separate memory traces to be encoded for repeated events. *Journal of Memory and Language*, *101*, 153-161.

Sahakyan, L., Delaney, P. F., & Waldum, E. R. (2008). Intentional forgetting is easier after two "shots" than one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 408–414. https://doi.org/10.1037/0278-7393.34.2.408

Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research: A theoretical and methodological review. *Psychology of learning and motivation*, *59*, 131-189.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145–166. https://doi.org/10.3758/BF03209391

Simmons, Joseph P. and Nelson, Leif D. and Simonsohn, Uri, A 21 Word Solution (October 14, 2012). *Available at SSRN:* https://ssrn.com/abstract=2160588 or http://dx.doi.org/10.2139/ssrn.2160588

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Smith, S. M., Handy, J. D., Hernandez, A., & Jacoby, L. L. (2018). Context specificity of automatic influences of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(10), 1501–1513. https://doi.org/10.1037/xlm0000523

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1589–1625. https://doi.org/10.1037/0278-7393.26.6.1589

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 529–536. https://doi.org/10.1016/0022-5371(76)90047-5

Toppino, T. C., & Gerbier, E. (2014). About practice: repetition, spacing, and abstraction. In B. H. Ross (Ed.), *The psychology of learning and motivation: Vol. 60. The psychology of learning and motivation* (pp. 113-189). San Diego, CA: Elsevier Academic Press.

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*, 1–12. https://doi.org/10.1037/h0080017

Unsworth, N., & Brewer, G. A. (2009). Examining the relationships among item recognition, source recognition, and recall from an individual differences perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1578–1585. https://doi.org/10.1037/a0017255

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory & Cognition*, *41*(1), 1-15.

Wahlheim, C. N., & Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General*, *148*(1), 30.

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 94–105. https://doi.org/10.1037/a0034055

Wahlheim, C. N., Smith, W. G., & Delaney, P. F. (2019). Reminders can enhance or impair episodic memory updating: A memory-for-change perspective. *Memory*, *27*(6), 849–867. https://doi.org/10.1080/09658211.2019.1582677

Whiffen, J. W., & Karpicke, J. D. (2017). The role of episodic context in retrieval practice effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(7), 1036–1046. https://doi.org/10.1037/xlm0000379

Wixted, J., & Mcdowell, J. (1989). Contributions to the Functional Analysis of Single-Trial Free Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 685–697. https://doi.org/10.1037/0278-7393.15.4.685

## APPENDIX A: RESEARCH ASSISTANT SCRIPT

**Study phase:** "You will soon view a list of words please try to remember these words for a later memory test."

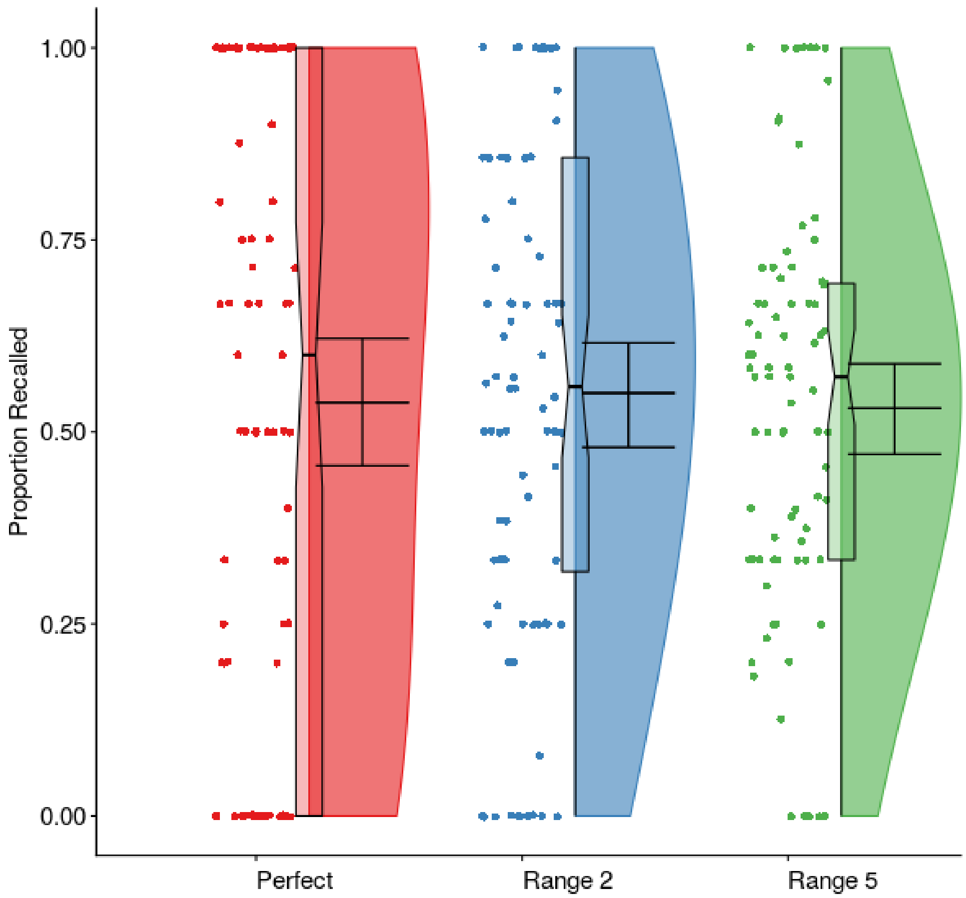
**Restudy phase:** "You will soon view a list of words. For each word please indicate if the item is new or if it has been previously shown. You will have the option to check "New or Old" after the item has been presented. If the item is old you will also indicate via slider where in the previous list you believe the item was. A check box will also be provided if you have no idea where in the list the item occurred. Points will be awarded for correctly stating if the item was new or old and additional points can be awarded for correctly stating the item's position in the list. A quarter point can be earned by checking the "no idea" option."

**Testing phase: "**For the final task you will have four minutes to recall as many words as you can from any part of the experiment. You will be awarded points for each correct answer."

# APPENDIX B: EXAMPLE OF SLIDER

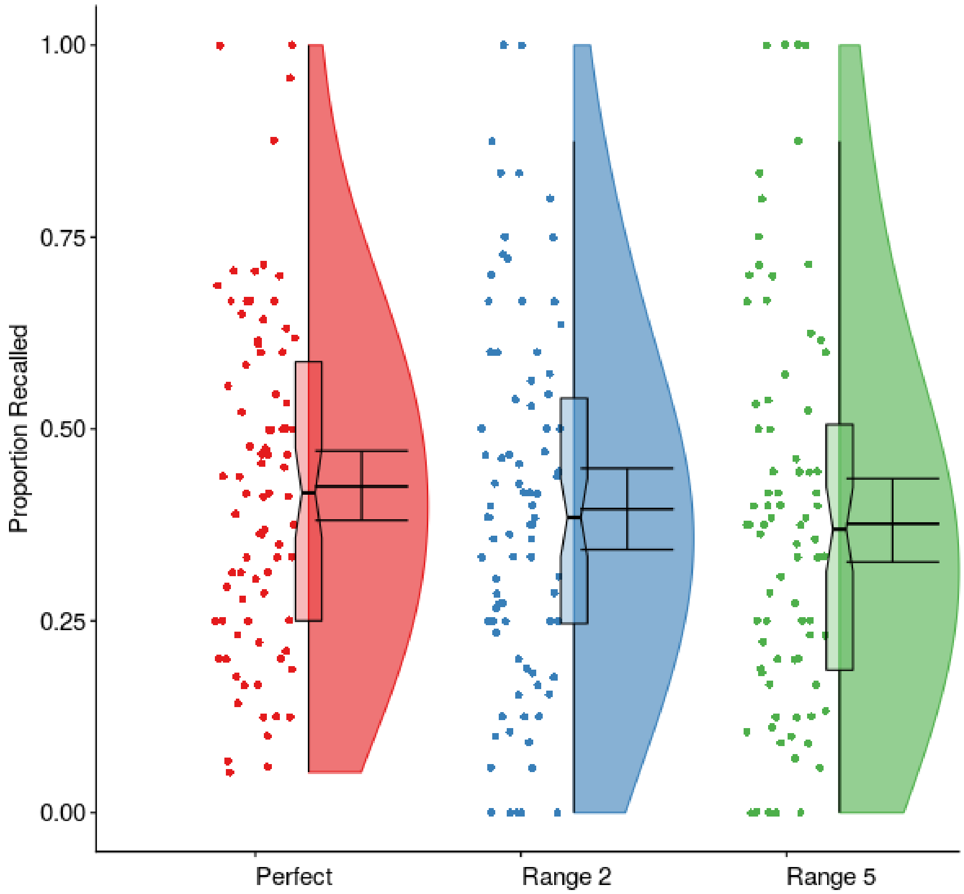

.

# APPENDIX C: EFFECTS OF ACCURATE MEMORY FOR WHEN AT REPETITION ON FINAL RECALL FOR DIFFERENT RANGES E1



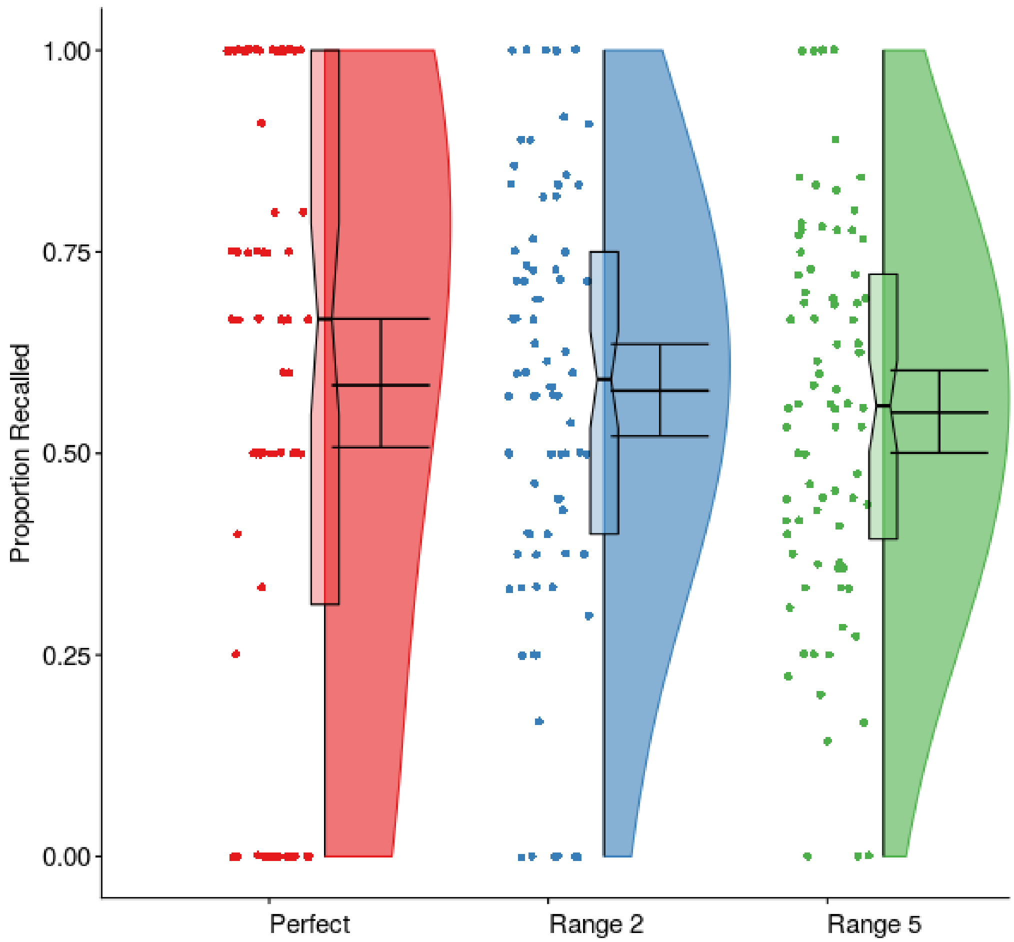

*Note.* Raincloud plot of the proportions of items recognized as old with MFW at three different tolerances during the restudy phase on final recall.

## APPENDIX D: EFFECTS OF INACCURATE MEMORY FOR WHEN AT REPETITION ON FINAL RECALL FOR DIFFERENT RANGES E1



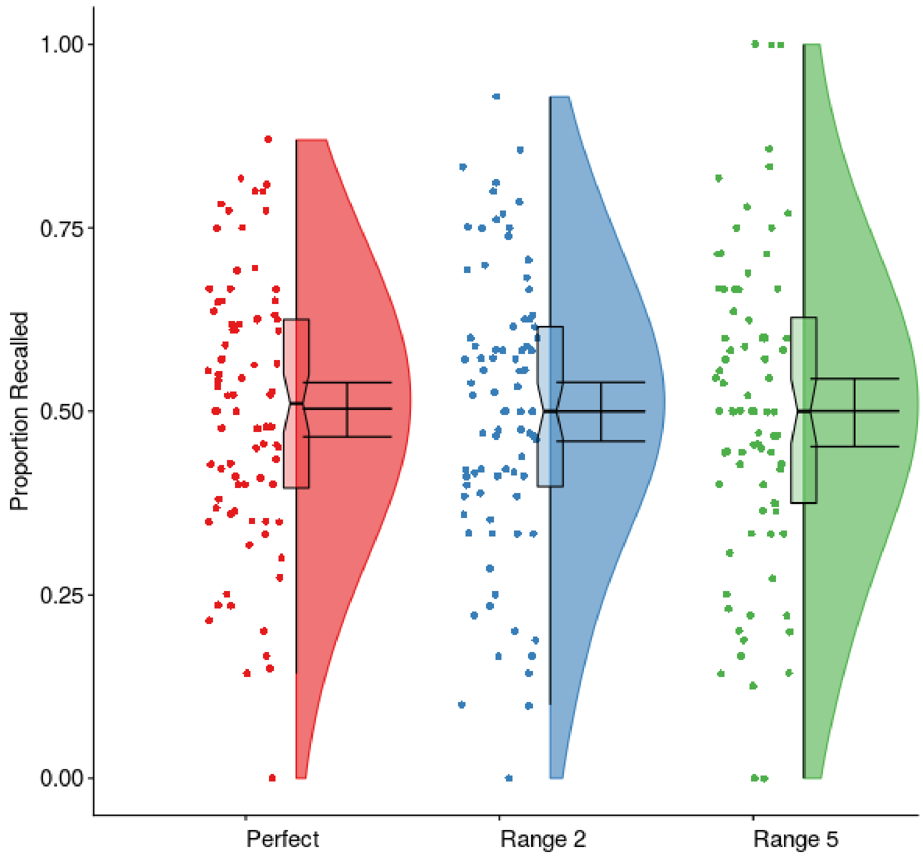

*Note.* Raincloud plot of the proportions of items recognized as old without MFW at three different tolerances during the restudy phase on final recall.

APPENDIX E: EFFECTS OF ACCURATE MEMORY FOR WHEN AT REPETITION ON FINAL RECALL FOR DIFFERENT RANGES E2




*Note.* Raincloud plot of the proportions of items recognized as old with MFW at three different tolerances during the restudy phase on final recall.

APPENDIX F: EFFECTS OF INACCURATE MEMORY FOR WHEN AT REPETITION ON FINAL RECALL FOR DIFFERENT RANGES E2




*Note.* Raincloud plot of the proportions of items recognized as old without MFW at three different tolerances during the restudy phase on final recall.

APPENDIX G: PROPORTION RECALLED MEANS AND STANDARD DEVIATION FOR RECALL CONDITION E1

| **Condition** | **M** | **SD** |
|---|---|---|
| Spaced | 0.40 | 0.21 |
| Singleton | 0.28 | 0.18 |
| Hits with MFW | 0.55 | 0.32 |
| Hits without MFW | 0.39 | 0.24 |
| Missed | 0.19 | 0.27 |

*Note.* Results for Experiment 1.

APPENDIX H: PROPORTION RECALLED MEANS AND STANDARD DEVIATION FOR RECALL CONDITION E2

| **Condition** | **M** | **SD** |
|---|---|---|
| Spaced | 0.52 | 0.18 |
| Singleton | 0.26 | 0.13 |
| Hits with MFW | 0.58 | 0.26 |
| Hits without MFW | 0.50 | 0.19 |
| Missed | 0.13 | 0.27 |

*Note.* Results for Experiment 2.