SMITH, JAMES RAY, M.A. The Effect of Context-Change on Retrospective Time Estimates. (2014)
Directed by Dr. Lili Sahakyan. 45 pp.

Research using the *list before last* paradigm demonstrates that retrieval of a previously learned list (L0) between encoding of two other lists (L1 and L2) leads to forgetting of L1 -- an effect attributed to internal context-change (e.g., Jang & Huber, 2008; Sahakyan & Hendricks, 2012). In the current studies, I manipulated the nature of the material participants were asked to retrieve between L1 and L2 (Experiment 1), or the temporal distance between L0 and the remaining lists (Experiment 2) in order to examine whether such manipulations induce greater levels of context change. Specifically, in Experiment 1, some participants retrieved a previously encoded list of words (L0), whereas others retrieved a short video that they were exposed to earlier. In Experiment 2, L0 was studied either 1-hour or 14 days prior to the session where L1 and L2 were studied. To assess the magnitude of context change, both L1 recall accuracy and retrospective time estimates were analyzed.

THE EFFECT OF CONTEXT-CHANGE ON RETROSPECTIVE TIME ESTIMATES

by

James Ray Smith

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
2014

Approved by

____
Committee Chair

APPROVAL PAGE

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

Committee Chair____________________________________
Lili Sahakyan

Committee Members____________________________________
Peter F. Delaney

____
Stuart Marcovitch

7/9/2014______________________
Date of Acceptance by Committee

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Imagine that you are at a conference, where you have just been introduced to a new person and you want to be sure that you remember his/her name. Like many, you may employ a common strategy – once you are told the person's name, you might begin your next sentence with the name (e.g., "James, nice to meet you"). This can be an effective strategy. By generating/retrieving the person's name, you help to ensure that the name has been properly encoded. However, imagine that the person then asks you a question about some research you completed years ago. After attempting to answer the question, you might find yourself in an embarrassing but familiar situation. You might realize that you no longer remember the person's name. This is not too surprising considering what is known about the effects of retrieval on forgetting.

Research shows that retrieval of some information can lead to forgetting of other information that presides in memory. In the laboratory, such effects are demonstrated by having participants study list(s) of items, then retrieving some of the items (or some of the lists), and observing the effects of such retrieval on the remaining items/lists that were not retrieved. Forgetting can take place in form of impaired access to an entire list of items, as shown in research involving the *list-before-last* (LBL) *paradigm* (e.g., Jang & Huber, 2008; Sahakyan & Hendricks, 2012; Shiffrin, 1970), but it can occur at the level of semantic representations of individual items, as shown in research using the *retrieval-practice paradigm* (e.g., Anderson, Bjork, & Bjork, 1994). Current research utilizes the LBL paradigm, whereby participants encode several lists, and after study of each list, they are told to retrieve the previous list (not the most recent). Prior work shows that the act of retrieving a previously encoded list of words causes forgetting of a more-recently learned list (e.g., Sahakyan & Hendricks, 2012; Sahakyan & Smith, 2014).

The aim of the current research is to address discrepant findings between two recently conducted investigations. Briefly, Delaney, Sahakyan, Kelley, and Zimmerman (2010) found that retrieval of events from longer ago caused greater forgetting of more recently encoded information, whereas Sahakyan and Hendricks (2012) found that regardless of the age of the retrieved memory, the magnitude of forgetting for more recent information is unaffected. The two studies utilized different methods and paradigms, with Sahakyan and Hendricks (2012) using the LBL paradigm, and Delaney et al. (2010) utilizing the diversionary thought paradigm (described later). These two paradigms are related by the nature of their underlying mechanism as both take their roots in research on directed forgetting. I describe each paradigm and related research in the next sections.

List-before-last Paradigm

The list-before-last (LBL) paradigm was originally invented by Shiffrin (1970) in an attempt to study whether forgetting in long-term memory was due to time decay or interference. In the classic version of LBL, participants study multiple lists, and starting with the second list, they are asked to retrieve the previous list rather than the current list. However, some studies have used an abbreviated version of the paradigm, whereby participants study only three lists (L0, L1, and L2, respectively) and, after L1 encoding, they are told to either retrieve L0 or to solve math problems as a control condition (e.g., Sahakyan & Hendricks, 2012). Afterwards, all participants encode L2 and are subsequently tested for their memory of L1 using a free recall test. The typical findings show that participants who are required to retrieve L0 in between L1 and L2 exhibit forgetting of L1 on the final test (e.g., Sahakyan & Hendricks, 2012; Sahakyan & Smith, 2013). In addition to the negative effects of retrieval, there are also positive effects of retrieval, whereby the interference from L2 on L1 is reduced. These positive effects of retrieval have been demonstrated in terms of reduced effects of list-length manipulation (e.g., Jang & Huber, 2008; Shiffrin, 1970) or in terms of reduced intrusions from L2 on L1 (e.g., Sahakyan & Hendricks, 2012). These effects have been explained in terms of L0 retrieval trials producing an accelerated mental context change between L1 and L2 (Jang & Huber, 2008; Sahakyan & Hendricks, 2012). *Mental context* can be broadly defined as any information that is actively represented in mind as people encode lists of items. It refers to incidental or background stimuli that are also present during encoding of a target item, including the spatial-temporal environment in which items are encoded, the internal states of the participants, as well as the thoughts they experience in response to the items during encoding. Some of these stimuli are perceived by the observer, whereas others may escape a person's deliberate attention. Nonetheless, these stimuli—whether intentionally processed or not— can serve as cues during the retrieval of target items and these context cues gradually fluctuate over time as new information is processed and as new cues are generated. Thus, over time, a mismatch begins to form between the current context and the context in which a previous item was learned. This mismatch between encoding and retrieval contexts—or context change—leads to reduced access to the encoding context thereby producing forgetting of those items (Estes, 1955; Mensink & Raaijmakers, 1988; Sahakyan & Kelley, 2002).

Sahakyan and Hendricks (2012) argued that in the math group of the LBL paradigm, the context associated with each list drifts from one list to the next list in a somewhat gradual fashion, thus the contexts between any two adjacent lists are more similar to each other than the contexts between more distally separated lists. However, in the retrieval group, retrieval of L0 disrupts the similarity of the contexts of the L1 and L2 lists, creating an accelerated contextual drift between them, as though by temporally shifting or segregating them from each other. During the final test, it becomes more difficult to reinstate the context of L1 due to a larger mismatch between L1 and L2 contexts. As a result, the L0 retrieval group recalls fewer L1 items than the math group. Thus, via internal context change, retrieval of a previous list causes the forgetting of the more-recently learned list (Jang & Huber, 2008; Sahakyan & Kelley, 2012, Sahakyan & Smith, 2014).

More relevant to the current investigation is one of the experiments of Sahakyan and Hendricks (2012), where the authors attempted to vary the degree of internal context change experienced by participants by varying the lag between L0 and the subsequent two lists (e.g., L1 and L2). Specifically, in an initial session, they had participants encode L0 and then asked them to return for a second session at a later time. This second session occurred one hour, 24 hours, or 72 hours after the first session. This meant that, for some participants, L0 was more temporally distant from L1 and L2 than for other participants. Sahakyan and Hendricks expected that an increase in the temporal distance of L0 would initiate greater levels of mental context change when L0 is retrieved between the remaining two lists, and that greater degree of mental context change should be evidenced by poorer recall of L1. However, their results did not support this prediction. They found that all groups that engaged in L0 retrieval experienced similar magnitude of L1 forgetting at the time of final test regardless of how long ago L0 was studied. In other words, retrieval of more temporally distant events from memory did not affect the degree of forgetting of more recently encoded materials.

These findings counter those of Delaney et al. (2010), who used the *diversionary thought paradigm* (described next), and found that retrieving more temporally distant events from memory caused greater forgetting of previously-learned list (compared to retrieving temporally-nearby events).

The Diversionary Thought Paradigm

The idea that L0 retrieval causes a larger shift in context between L1 and L2 is derived from previous research on *directed forgetting* (DF) and the related diversionary thought experiments (e.g., Sahakyan & Kelley, 2002). In a typical list-method directed forgetting (DF) experiment, participants are given two lists of words (L1 and L2). In between the lists, some participants are told to forget the first list whereas others are told to remember the first list. The results of such experiments indicate that participants in the forget condition are indeed able to forget L1 items when told to do so compared to the remember group. Sahakyan and Kelley (2002) collected retrospective verbal reports from their participants, many of who reported that in response to the forget instruction, they tried to think of other things unrelated to the experiment in order to forget unwanted items. Building upon such evidence, Sahakyan and Kelley (2002) showed forgetting of L1 in the absence of any specific instructions to forget. By employing what has become known in the literature as the *diversionary thought paradigm*, they asked participants to think of something else pre-specified by the experimenter in between the study of two lists of words (e.g., "imagine if you were invisible"). When L1 memory was tested during the final test, participants in the diversionary thought condition (similar to those who are told to forget L1) showed comparably poorer recall compared to the control participants who did not engage in diversionary thought and solved math problems instead. Sahakyan and Kelley (2002) proposed that engaging in distracting thoughts shifts participants' mental context between the two lists, making it more difficult to reinstate the original context of the to-be-forgotten list, thus creating impaired recall of L1.

The diversionary thought paradigm has become a well-validated method of inducing internal context change (e.g., Delaney et al., 2010; Pastotter & Bauml, 2007; Unsworth, Spillers, & Brewer, 2011). The type of diversionary thoughts that participants were asked to engage in varied from study to study. For example, participants have been asked to imagine being invisible (Sahakyan & Kelley, 2002), imagine their childhood home (Sahakyan & Delaney, 2003), daydream about vacations (Delaney et al., 2010), or even engage in a task that does not engage diversionary thought per se but more like a diversionary action such as to wipe a computer monitor (Mulji & Bodner, 2011). The results across all these studies mimic those of a DF experiment. When L1 memory is tested during the final test, participants in the diversionary thought/action condition—similar to those who are told to forget—show poorer recall compared to the control participants.

In one of the diversionary thought experiments, Delaney et al. (2010) had participants study two lists and, in between those lists, asked participants to retrieve memories about their current home or about their childhood/parents' home. They found that participants experienced more L1 forgetting at the time of final test as a consequence of retrieving memories related to their parents' home compared to memories of their own home. More importantly, the longer it had been since participants had visited their parent's home, the more L1 forgetting they exhibited. In other words, if the to-be-retrieved memory was from a more temporally distant past, it created more forgetting of the current list (e.g., L1). Delaney et al. (2010) attributed greater forgetting of L1 to larger change of mental context taking place between L1 and L2 as a consequence of retrieving *more temporally distant* events.

Thus far, one study showed that retrieval of more distant events causes more forgetting of the current event (e.g., Delaney et al. 2010), whereas the other study showed that retrieval of more distant events causes the same magnitude of forgetting (e.g., Sahakyan & Hendricks, 2012). Given that there were several major methodological differences between the two studies, the current investigation aimed to uncover the factors that could contribute to the discrepancy in findings.

Assessing Context-Change through Retrospective Time Estimates

Although mental context-change is an important theoretical construct central to formal theories of memory, it is typically inferred from assessing various indices of memory performance, including accuracy measures, response latency measures, as well as analyzing the dynamics of retrieval process itself, such as where in the list participants initiate retrieval, and how they transition between responses during retrieval. Overall, internal context is hard to experimentally manipulate in the lab, because it is difficult to define independent of memory performance. Our recent research suggests that retrospective duration estimates could be used as independent indicators of internal context change (Sahakyan & Smith, 2014).

Retrospective time judgments are those made when participants are unaware that they will be asked to make a time estimate until after the time period has already passed. Because they are unaware that such information will be assessed, it is unlikely that participants will consciously attend to temporal information during the target interval, and hence time estimates must be based on information available from memory (for a review, see Block & Zakay, 1997). Prior research shows that the type of information that people use in making these estimates involves memory for changes in context. For example, when two lists are studied in different physical contexts as opposed to the same context, the retrospective time estimates are longer in the different context condition compared to the same context (e.g., Block, 1982). Likewise, changes in cognitive context that result from encoding the items using different orienting tasks throughout the list as opposed to the same orienting task also lead to overestimation of the target interval (Block & Reed, 1978). Along the same lines, of the two intervals of the same objective duration and similar information processed, a segmented interval is remembered as being longer than an unsegmented interval (Poynter, 1983; Zakay, Tsal, Moses, & Shahar, 1994).

Finally, our own research using memory paradigms that are typically interpreted in terms of internal context-change, such as the DF paradigm (Sahakyan & Kelley, 2002), or the LBL paradigm (Jang & Huber, 2008) confirms that participants in conditions attributed to internal context-change provide significantly longer time estimates compared to participants in the respective control conditions (Sahakyan & Smith, 2014). Therefore, in the current experiments, I collected retrospective times estimates as additional measures of context-change to obtain corroborating evidence of context-change that was independent of memory assessment.

CHAPTER II

EXPERIMENT 1

One of the critical methodological differences between the Delaney et al. (2010) study and the Sahakyan and Hendricks (2012) study is the nature of the event participants were asked to retrieve between L1 and L2. Namely, in the Delaney et al. (2010) study, participants were asked to recall their childhood home, whereas in Sahakyan and Hendricks (2012), participants were asked to recall a previously studied list of words. It is possible that retrieval of dissimilar material (e.g., childhood home) creates more forgetting than retrieval of similar material (e.g., another list) because when the mind jumps back in time to retrieve previously encoded material that is so different from the current materials, it changes the current state of context between L1 and L2 in a more accelerated fashion than when the mind jumps back to more similar material. The similarity between the materials that are presiding in memory and what needs to be retrieved may safeguard against rapid contextual drift, whereas retrieval of dissimilar materials may drift the mental context to a greater extent. To investigate these possibilities, I manipulated the nature of the material retrieved between the two lists. Some participants retrieved L0 between L1 and L2, whereas others retrieved a previously seen video.

By using a video segment, I was better able to control for the subjective nature of the previously-used diversionary thought tasks (e.g., think about childhood home). Asking participants to recall their parents' home may elicit a different experience for each participant as no two homes are the same—in terms of temporal distance, physical features, and even personal salience. Previous researchers (Smith & Manzano, 2010) have utilized video in context-change research with success, and a video stands as a much more controlled and consistent exercise in diversionary thought. If dissimilarity of materials is responsible for discrepant findings between the Delaney et al. (2010) and Sahakyan and Hendricks (2012) study, I expected that recalling a video between L1 and L2 would lead to the greater forgetting than recalling previously encoded L0. In addition, if retrieval of video creates greater internal context-change between L1 and L2, then I expected to obtain greater overestimation bias in time estimates in the Video group compared to the List group.

CHAPTER III

METHODS

Participants

A total of 85 participants were recruited in this experiment from the participant-pool at the University of North Carolina at Greensboro. The experiment was completed in exchange for course credit.

Materials

Three lists (L0, L1, and L2) of unrelated words were used (see Appendix A). Each list contained 12 medium-frequency, unrelated nouns. The words were five or six letters long, presented in random order, and the presentation order of the lists was fully counterbalanced. Each list consisted of six animate and six inanimate objects because, as part of one of the orienting tasks, participants were asked to judge the words as animate or inanimate during encoding. The video used in Session 1 was four minutes and 40 seconds long, and consisted of a brief primer on perceptual illusions. The video was colorful and engaging to ensure that participants paid close attention to the video and had plenty of possible information to encode. The video can be found at http://www.youtube.com/watch?v=PN1NAiM55hU&feature=related. Experimenters were provided digital watches in order to keep track of elapsed time when appropriate.

The music played in Session 1 was the “Star Wars Theme” played through the same computer speakers as the experiment.

Design

Each participant completed two separate sessions (Session 1 and Session 2). In Session 1, all participants were shown a list of words (L0) and also a brief video (in a counterbalanced order). In Session 2, which took place an hour later, all participants encoded two lists of words (L1 and L2), and performed an intervening task between the two lists. There were three groups in the experiment based on the type of task completed between the lists in Session 2. Some participants recalled L0 from Session 1 (termed the List group), some recalled a video from Session 1 (the Video group), and the remaining participants did not engage in retrieval of previously encoded event, but instead solved arithmetic problems (the Math group). The critical variables of interest involve memory for L1 as well as the subjective time estimates for Session 2.

Procedure

Before each session began, participants were asked to store their belongings in a separate and locked room. This ensured that cell phones were not a distraction during any part of the experiment and this also denied participants the ability to check their phone for the time. The main theme from Star Wars was played as participants entered the experiment room. This music was included in order to mark the experiment as a unique event for participants (important for later recall of the episode). In addition, I wanted to elicit encoding of contextual information in more than one modality. The experimenter shut off the music after the participant had read and signed the consent form and, at this point, the experimenter began the instructions. The instructions were presented in pre-recorded audio format in order to equate the duration of the experiment across participants. Care was taken to ensure that the instructions were presented through the computer speakers at a specific, pre-set volume for each trial. This volume corresponded to the seventh setting on the speaker volume dial (Dell-Altec Lansing speakers model number CN-0W2739) while all computer settings were set at the maximum possible volume. No participants reported difficulty in hearing the instructions.

Because participants left the lab for one hour between Session 1 and Session 2, to prevent rehearsal between the sessions, an incidental memory task was used in Session 1. Participants were told that they were helping the experimenter choose a list of words for use in future experiments. During Session 1, all participants were presented with List 0 as well as an instructional video, and the order in which they were presented was counterbalanced—some participants saw the list first (three times), whereas others saw the video first (presented once). List 0 was shown three times in different random orders each time and, through each presentation, participants were asked to perform a different orientating task. All responses were made using a computer keyboard. Multiple presentations of L0 were used to safeguard against floor recall levels given that Session 2 took place after a delay. Throughout the first and the second presentation of L0, the words were presented at a rate of four seconds per item with a one second inter-stimulus interval (ISI). During the first presentation of L0, participants were asked to provide a pleasantness rating for the words on a scale from 1-7, whereas during the second presentation, they were asked to categorize the word as either animate or inanimate. During the third presentation, they were asked to make up a sentence using the word and type it. During this third presentation of the list, the presentation rate was self-paced. Participants were given as much time as needed to create each sentence, and they pressed the space-bar to advance to the next word.

After L0 was presented three times, participants were given a 90 second free-recall test for all L0 words they could remember. An immediate test was also used by Sahakyan and Hendricks (2012) to minimize potential floor effects after a delay given that retrieval safeguards against forgetting (e.g., Roediger & Karpicke, 2006). The free recall test was carried out on the computer, and the screen displayed each response so that participants could see all the words they recalled. After L0 encoding and test, participants were also shown an instructional video that lasted 4 minutes and 40 seconds (or vice versa – they viewed video first, followed by L0 encoding/test).

For the video portion, participants were told they would be evaluating videos for use in freshman-level psychology courses. After viewing the video, they were asked to provide three ratings for the video on a 7-point Likert scale — (1) the effectiveness of the video at explaining the material, (2) the difficulty of the material being discussed, and (3) the overall quality of the video. These ratings were made by typing responses into a computer. After rating the video, participants were told to return for Session 2, which took place one hour later. No details were provided as to the nature of Session 2.

Consistent with previous research, Session 2 responses in all conditions were made on a separate sheet of paper instead of typing them on the computer screen (Sahakyan & Hendricks, 2012; Unsworth, Spillers, & Brewer, 2012). During Session 2, participants were told that they would study two lists of words (L1 and L2), and that they should remember them for an upcoming memory test. Thus, encoding in Session 2 was intentional rather than incidental consistent with previous research with the LBL paradigm. After the instructions, a red screen with a brief tone lasting two seconds appeared. Participants were told to expect this screen with a tone as it marked the moment that the experiment officially began. I marked the start of the experiment with the tone so that later on I could redirect participants' attention to a defined "start" of the session, and ask them to provide a retrospective estimate of how much time has passed in the experiment since that moment (that is, estimate the duration of Session 2). After the red screen and the tone, participants were then oriented to the center of the computer screen with a plus sign and then shown L1 words. The words in L1 were also presented at a rate of four seconds per word with a one second ISI. There was no orienting task for either L1 or L2. After L1 encoding, some of the participants were asked to solve math problems for 90 seconds while the remaining participants were asked to either recall L0 from Session 1 for 90 seconds, or to retrieve the video they were shown in Session 1 for 90 seconds. Specifically, participants asked to recall the video were told to mentally reinstate their experience of the video and were prompted not only to think of the factual elements of the video, but to also imagine any sounds, colors, and thoughts they may have experienced during the entire video.

After solving math, or retrieving L0 or the video, all participants were instructed to study L2 just as they had studied L1. Once again, the words were presented for four seconds each with a one second ISI. The last word of L2 was followed by another red screen and tone (the same tone used previously) so that the end of the study phase of the experiment was also clearly marked for the participants. After the tone, all participants were asked to estimate how much time had passed between the two tones. Participants wearing a watch were not told to remove it. This is a measured decision based on a meta-analysis of time judgment research that found that retrospective judgments tend to be significantly longer for participants who are told to remove their watch compared to those that are told nothing presumably because it draws attention to time, which nullifies the incidental nature of the task (Block & Zakay, 1997). Experimenters were instructed to monitor whether participants explicitly checked the time but, given that cell phones are now the primary source for this information, the presence of watches was rare. Nonetheless, zero instances were reported by experimenters.

A space was provided on the computer screen for participants to type time estimates. They were asked to give their estimate as accurately as possible—on the order of seconds. Finally, participants were given 90 seconds to recall L1 by typing the words.

CHAPTER IV

RESULTS

The probability of a Type 1 error was set to .05 for all analyses. First I report memory accuracy to assess the magnitude of L1 forgetting effects, and then I report the time estimation data. As is conventional with time estimation research (Roy & Christenfeld, 2007, 2008), the outliers above 2.5 standard deviations from the grand mean for time estimation were removed from the analyses. Given the hypothesized relationship between time estimates and recall, this was also done for L1 recall results (totaling one participant for time estimation results and zero for recall results).

Recall Results

Given the incidental nature of the Session 1 encoding tasks, I examined overall recall rates for Session 1 to assess the degree of learning. The Session 1 average recall proportion for L0 was quite high (M = .72, SD = .12). After the one hour delay, during Session 2, participants' recall of L0 (in between L1 and l2) was still relatively high (M = .55, SD = .14). My critical prediction concerned L1 recall, and thus I assessed proportion of L1 recalled using a one-way ANOVA with Group (Video vs. List vs. Math) as a factor. The results are shown in Figure 1. There were no significant group differences, $F(2,81) = 1.23$, $p = .30$, $\eta^2 = .03$

That is, regardless of the task performed between L1 and L2, the differences in L1 recall did not reach statistical significance. Further analyses also failed to support my initial prediction. It was expected that the Video group would recall fewer L1 words compared to the Math group (the control group). Planned comparison between the Video and the Math group revealed that the difference in L1 recall did not reach significance although numerically it was in the predicted direction, $t(55) = 1.41$, $p = .16$. The same pattern was obtained also between the List group and the Math group, although numerically the difference was again in the predicted direction, with the List group recalling fewer words than the Math group, $t(57) = 1.45$, $p = .15$. Thus, although the group differences did not reach statistical significance, numerically they followed the predicted trend with participants who recalled the video or the list performing the worst, followed by participants in the math control group who recalled the most L1 words. These results cautiously suggest that a similar magnitude of internal context change may have taken place in the Video and the List group, but given that the differences were not statistically significant, evidence from additional dependent measures is warranted to support these tentative conclusions.

Figure 1. Proportion of L1 Recall in Experiment 1

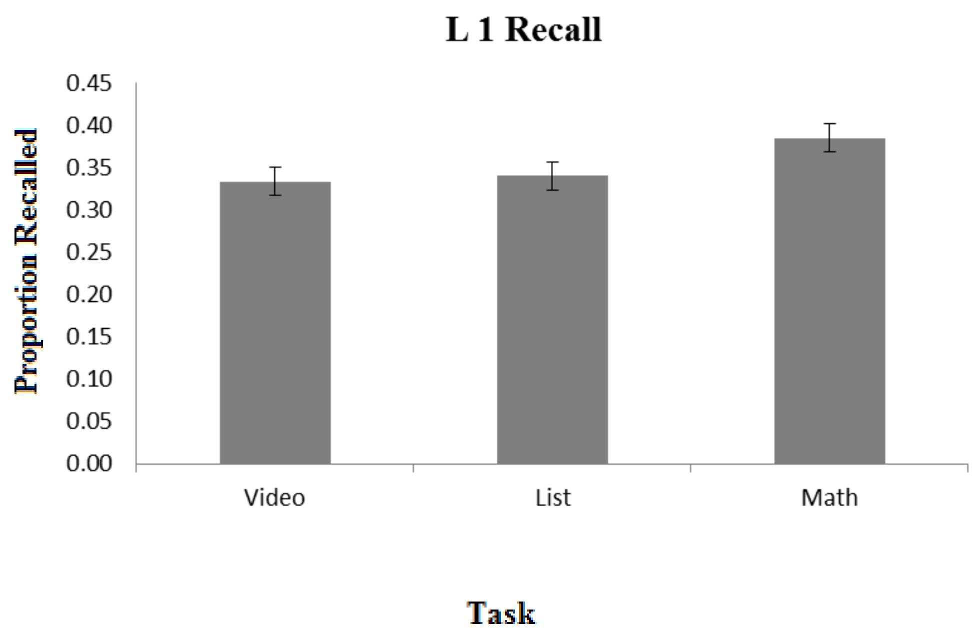

Figure 1. Proportion of L1 recalled as a function of task/condition. The error bars represent ±*SEs* of the mean.

Time Estimate Results

In prior work, we found that participants experiencing internal context-change were greatly overestimating the duration of the experiment (Sahakyan & Smith, 2014). Hence time estimates could provide an independent way of assessing the magnitude of context-change in addition to memory. Thus, if the two retrieval groups (List and Video) experienced internal context change, then I would expect them to overestimate the subjective duration of the experiment compared the math control group.

The duration of Session 2 was exactly 5.5 minutes. To determine if there is overestimation or underestimation bias, it is typical to calculate the ratio of subjective duration to objective duration, where positive values indicate overestimation, and negative values indicate underestimation (e.g., Block & Zakay, 1997). Block and Zakay (1997) also point out that these duration ratios allow for comparisons across conditions and experiments that utilize different durations—an important step in standardizing measures across all time estimation research. Finally, the recommended procedure is to log-transform duration ratios because underestimation is bounded by the lower limit because the subjective time estimates cannot be less than zero, whereas overestimation is unbounded by the upper limit given that theoretically subjective estimates can be any number (Roy & Christenfeld, 2007, 2008).

A one-way ANOVA using Group as a factor was performed on Session 2 log-transformed duration ratios (average skewness = -0.08). The results are illustrated in Figure 2. The results failed to reach traditional significance, $F(2,81) = 2.24$, $p = .11$,

$\eta^2 = .05$. Comparing the duration ratios of the Video group to the Math control revealed significant difference between the two groups, $t(55)= 2.27$, $p < .05$. However, the difference between the List and Math groups did not reach significance, $t < 1$. If the two retrieval groups (Video and List) experienced similar magnitude of context change, they should overestimate the duration, and thus I performed a one-sample *t*-test comparing the duration ratios to zero (e.g., accurate estimate). This analysis revealed significant overestimation bias in the Video group, $t(24) = 2.10$, $p < .05$. However, the List group neither overestimated nor underestimated the duration as its duration ratio was not significantly different from zero, $t < 1$. The same was true also in the Math group, $t(31) = 1.57$, $p = .13$. Only the Video group's estimate was significantly different from zero, and a comparison of the time estimates for the List and Video groups did not yield significant results, $t(51) = 1.15$, $p = .26$.

Overall, the subjective estimates followed a somewhat graded pattern with the participants who recalled the video overestimating the duration of Session 2 the most, followed by those who recalled L0, followed by the math control group. However, only the Video group significantly overestimated the duration of the experiment, whereas the List group did not.

Figure 2. Time Estimation Accuracy for Session 2 in Experiment 1

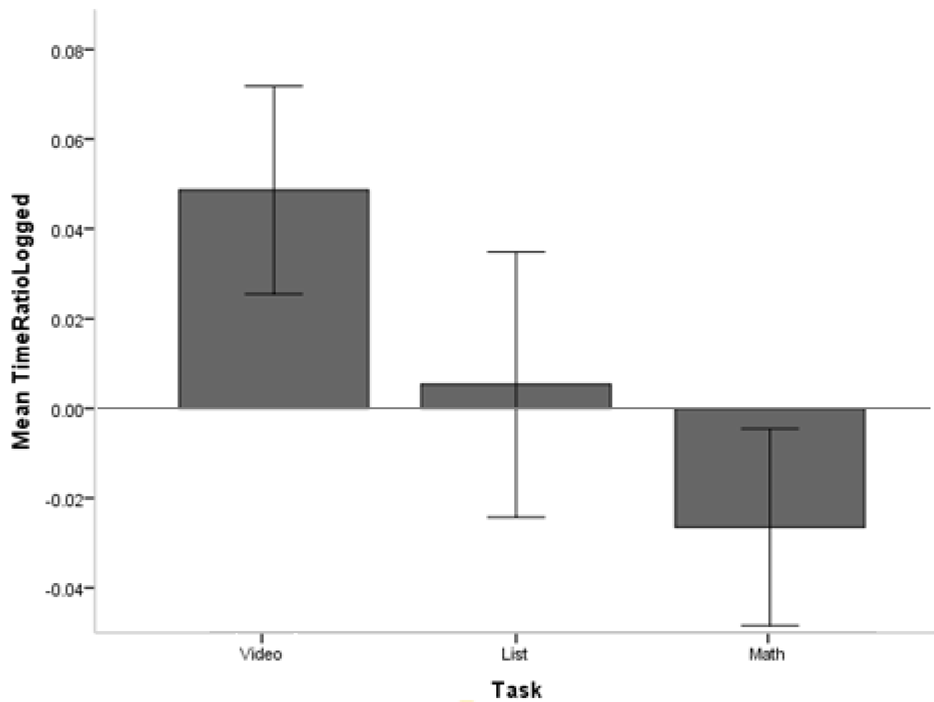

Figure 2. Log-transformed duration ratios for Session 2 in Experiment 1. The value 0 corresponds with an estimate equivalent to the objective duration. The error bars represent ±*SEs* of the means

Discussion

All in all, the results do not support my initial predictions. In fact, as traditionally measured by L1 recall, the manipulations used in Experiment 1 did not produce measureable and reliable contextual change for participants, as the differences in L1 recall rates for the Video and List groups did not differ significantly from the Math control. Furthermore, only the Video group significantly overestimated the Session 2 duration. Although the individual analyses of recall and time estimates did not offer evidence to support my hypotheses, taken together, they suggest that the type of information retrieved from memory may play a role in initiating context change. In prior research (e.g., Sahakyan & Smith, 2014), retrospective time estimates in the internal context-change conditions showed significant overestimation bias. Namely, conditions with most forgetting caused by internal context-change were associated with higher time estimates. If duration judgments are markers of internal context change, then one would expect an inverse relationship between the recall rates and time estimates as found by Sahakyan and Smith (2014). However, this was only partially the case in the current experiment. The L1 recall rate was lower in the Video group than in the Math group, and the Video group overestimated the duration of the experiment compared to Math control. The lower recall and higher time estimates in the Video group are consistent with the notion that this group likely experienced greater internal context-change compared to the Math group.

Interestingly, although L1 recall rates were similar for the List and Video groups, only the Video group overestimated the Session 2 duration. The List group recalled fewer L1 words than the Math group (although the difference was not statistically significant), but this group did not overestimate the duration of Session 2. The List group's time estimates were unbiased. Thus, it is possible that retrieving a video between L1 and L2 may be fundamentally different from recalling a list or doing basic math problems in terms of initiating a greater internal context change. While only the Video group's estimate was significantly different from zero, a comparison of the time estimates for the List and Video groups did not yield significant results. Given this overall pattern of results, it is also possible that the type of information retrieved between L1 and L2 may not influence context change as measured by L1 recall or Session 2 duration estimates—at least, not on this time scale (discussed later).

CHAPTER V

EXPERIMENT 2

As previously mentioned, Delaney et al. (2010) found that the magnitude of internal context change (as reflected in amount of L1 forgetting) depends on the temporal distance of an imagined event/place between the two lists. Namely, between L1 and L2, they had participants remember their parents' house (or their own house, in the control condition) and found that the longer it had been since participants had visited their parents' house, the fewer words they recalled from L1. Thus, the amount of time that has passed between the event and its retrieval time appears to be an important factor in driving internal context change. However, a different result was obtained by Sahakyan and Hendricks (2012), where between L1 and L2 participants were asked to retrieve a previously encoded list (i.e., L0), which was studied either on the same day, one day prior, or three days prior to the session where the remaining two lists were studied. Sahakyan and Hendricks (2012) found that although retrieving L0 produces internal context change and leads to L1 forgetting (compared to solving math between the lists), the magnitude of forgetting was unaffected by how long ago L0 was encoded.

The delay was varied over a period of a few days in Sahakyan and Hendricks (2012) study, whereas the delay in the Delaney et al. (2010) study involved several weeks. Delaney and colleagues report that, on average, participants had visited their parents' house 5.1 weeks ago, whereas, participants asked to imagine their own home reported having visited that home only 2.1 hours before (on average). The results of Delaney et al. (2010) showed that the delay on the order of hours was irrelevant, but that a delay on the order of weeks was more critical for producing differential degree of internal context change and concomitant forgetting. Therefore, the use of different delay intervals across the two studies may be one reason for discrepant findings.

In Experiment 1, I found that retrieving L0 that was studied on the same day produces as much forgetting of L1 as retrieving a video, suggesting that the nature of event retrieved might not matter for the magnitude of context change. At the same time, time estimates suggested that the video group experienced greater context change than the list group. It could be that a longer delay separating the sessions where L0 is studied and where L1 and L2 are studied is needed to obtain more conclusive evidence about the impact of the nature of retrieved events across both dependent measures.

For this reason, in Experiment 2, I increased the temporal distance between Session 1 and Session 2 to two weeks (instead of an hour, as in Experiment 1). Doing so allowed the retrieved event (e.g., L0) to fall more closely in line with the temporal distance of the events imagined in the Delaney et al. (2010) study. I expected that greater delay between Session 1 and Session 2 should create greater context change between L1 and L2, thus leading to greater forgetting of L1 during the final memory test in the list group. In addition, a longer delay might allow for a more sensitive comparison between the Video and the List group. Hence, if the nature of retrieved event influences the magnitude of context change, I might be more likely to detect significant difference between the Video group and the List group in Experiment 2 due to longer delay manipulation.

CHAPTER VI

METHODS

Participants

A total of 78 participants were recruited in this experiment from the participant-pool at the University of North Carolina at Greensboro. The experiment was completed in exchange for course credit. None of the participants took part in Experiment 1.

Materials

The materials, including words lists, were the same as those used in Experiment 1.

Procedure

The procedure was the same as in Experiment 1, with a few exceptions. First, the delay between Session 1 and Session 2 was lengthened from one hour to two weeks. Second, when participants were told to retrieve L0 between L1 and L2 encoding, they were provided with the first two letters of each L0 word. A change from free recall of L0 (in Experiment 1) to cued recall of L0 (in Experiment 2) was done in order to avoid potential floor effects after a two-week delay. Thus, participants were given a sheet of paper containing all twelve cues at once. Importantly, in some of their experiments, Sahakyan and Hendricks (2012) also used cued recall of L0 instead of the free recall of L0, and they found that both types of tests lead to L1 forgetting on the final test.

CHAPTER VII

RESULTS

The probability of a Type 1 error was set to .05 for all analyses. Once again, I report memory accuracy measures to assess the magnitude of L1 forgetting effects, then I report the time estimation analyses. Once again, the outliers above 2.5 standard deviations from the grand mean for both L1 recall and time estimation were removed from the analyses (totaling one participant for recall and one participant for time estimation results).

Recall Results

In Session 1, the average recall rate for L0 was $M = .73$, $SD = .11$. In Session 2, this time after a two-week delay, the average recall rate was lower than in Session 1, ($M = .23$, $SD = .17$), but not at floor.

A one-way ANOVA on the proportion of L1 recall using Group (List vs. Video vs. Math) as a factor revealed no significant differences, $F < 1$ (see Figure 3). Thus, no L1 recall differences were found based on the type of task completed between L1 and L2. The nearly identical recall rates suggest that either context change did not occur for any of the participants in this experiment, or alternatively, everybody experienced similar magnitude of context change. Analyses of time estimation are reported next as they might provide further insight about the nature of recall findings.

Figure 3. Proportion of L1 Recall in Experiment 2

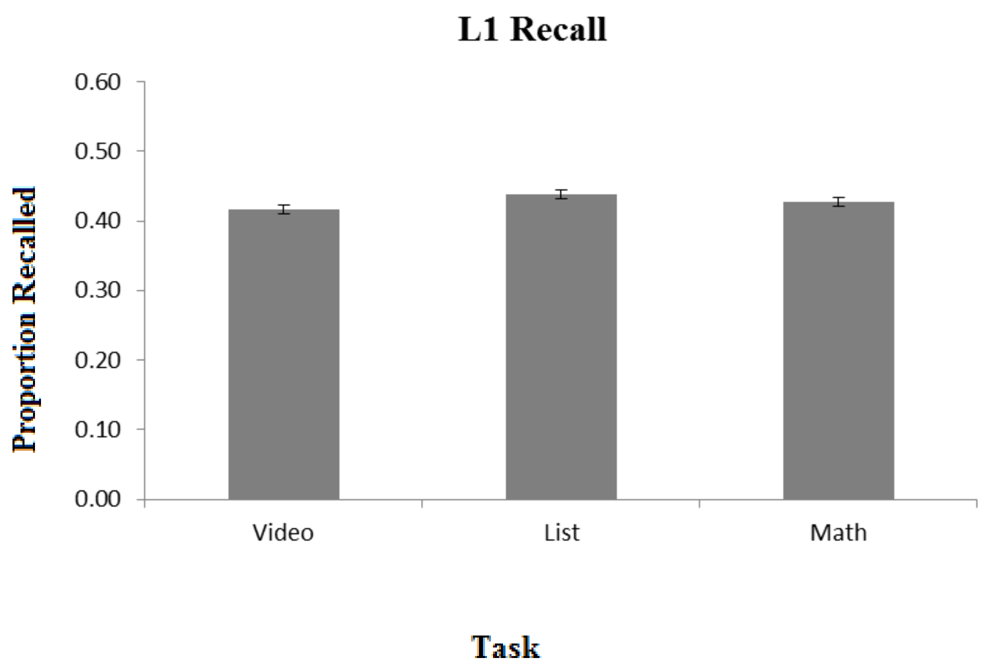

Figure 3. Proportion of L1 recalled as a function of task/condition in Experiment 2. The error bars represent ±*SEs* of the means.

Time Estimate Results

A one-way ANOVA using Group as a factor was performed on Session 2 log-transformed duration ratios (average skewness = -0.31). As depicted in Figure 4, there were no significant group differences, $F(2,73) = 1.03$, $p = .36$, $\eta_p^2 = .03$. Follow-up t-tests also revealed no significant group differences with neither the Math vs List comparison, $t(50) = 1.36$, $p = .18$, the List vs Video comparison, $t(49) = 1.25$, $p = .22$, nor the Math vs Video comparison, $t < 1$, reaching statistical significance. The Video group overestimated the duration numerically, but this overestimation was not significant, $t < 1$. This was also true for the Math group, which also overestimated the duration numerically, but the results were not significant, $t(24) = 1.00$, $p = .33$. Finally, the List group's estimate did not reveal any overestimation or underestimation biases, $t < 1$.

Figure 4. Time Estimation Accuracy for Session 2 in Experiment 2

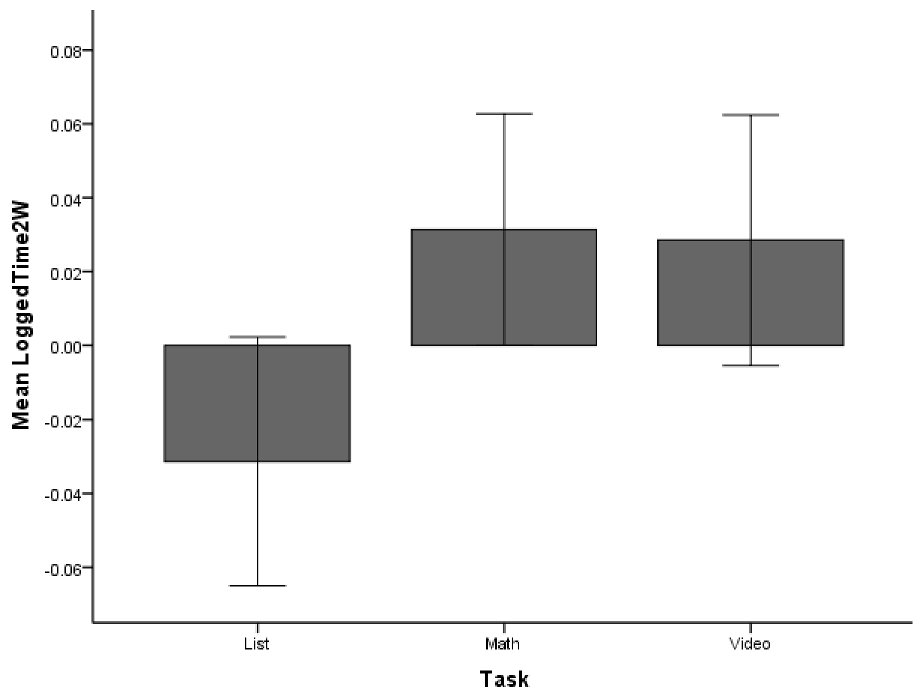

Figure 4. Log-transformed duration ratios for Session 2 in Experiment 2. The value 0 corresponds with an estimate equivalent to the objective duration. The error bars represent ±*SEs* of the means.

Discussion

Similar to the recall results, the results of time estimation data suggest that either no group experienced context change, or all groups experienced context change. Previous research shows that conditions experiencing context-change (reflected in lower accuracy in those conditions) overestimate the length of the interval. Thus, one would expect that if all groups experience context change, then the time estimates would be biased toward overestimation. However, in this experiment, the time estimation results did not differ significantly from zero, implying unbiased estimates (or low statistical power to detect such effects). Given that the recall and time estimation data for the List and Video conditions are equivalent to the Math control group, it is likely that neither experimental condition experienced context change. This conclusion is also supported by the fact that, although a cross-experiment comparison, the recall rates for all groups in Experiment 2 mirror the recall rates for the Math control group of Experiment 1.

CHAPTER VIII

GENERAL DISCUSSION

In two separate experiments, I attempted to expand on our previous work (Sahakyan & Smith, 2014), as well as resolve the discrepant findings of Delaney et al. (2010) and Sahakyan and Hendricks (2012). In Experiment 1, both the Video and List groups recalled fewer L1 words than the Math control group yet only the Video group overestimated the Session 2 duration. Furthermore, the differences in recall performance across the groups did not reach statistical significance. Despite this fact, the results hinted at a trend somewhat similar to the predicted outcome. The same cannot be said for Experiment 2, for which I introduced a longer delay between Session 1 and Session 2 (compared to Experiment 1). This second experiment yielded insignificant results for both dependent measures. In fact, there were no differences in recall performance or duration estimation across all groups

It was my expectation that the discrepancy in the results of Delaney et al. (2010) and Sahakyan and Hendricks (2012) was not only due to the nature of the items retrieved between L1 and L2, but also because of the difference in delays used across both studies. Based on previous research (Sahakyan & Smith, 2014), I predicted that conditions of context change would elicit poorer L1 recall and higher Session 2 duration estimates. Thus, I expected an inverse relationship between recall and duration estimates. I hypothesized that retrieving a video (which I anticipated would mimic retrieving a memory of a vacation/home) should initiate greater levels of context change compared to recalling a list or doing math problems, and I expected that this trend might be more pronounced if more time had elapsed between the encoding and retrieval of the target information. Nevertheless, the results of Experiment 1 and Experiment 2 do not support my predictions.

Clearly, more research is needed to determine the exact relationship between context change, time estimates, and recall. Given the results of the current investigation, it is not possible to determine the cause of the discrepancy between the Sahakyan and Hendricks (2012) and Delaney et al. (2010) studies. While the results of the two experiments do not offer conclusive evidence and they are difficult to reconcile with previous research, some factors and variables may help explain some of the discrepancies between the current findings and past research.

For example, in Experiment 1, the overall ANOVA performed on L1 recall and Session 2 time estimates did not yield significance, although the individual t-tests provided some evidence of group differences, and the graded pattern of both recall and time estimation results resembled the predicted pattern to some extent. Post hoc power analyses suggest that these null results may be due to a lack of statistical power. A meta-analysis by Smith and Vela (2001) reports that the average effect size for a context change experiment involving free recall is *Cohen's d* = .29. A post-hoc power analysis reveals that Experiment 1 had only a power of .64 to detect an effect this size. Similarly, Block and Zakay (1997) performed a meta-analysis on experiments involving retrospective time estimates and found an average effect size of *Cohen's d* = .33. A post-hoc power analysis reveals that Experiment 1 only had a power of .76 to detect such an effect. Although these analyses are post-hoc, they suggest that Experiment 1 was possibly underpowered.

In Experiment 2, because of the long delay between Session 1 and Session 2, the List and Video groups may not have been able to remember the target information from Session 1. The two-week delay mimics the delays found in Delaney et al. (2010), but recalling a vivid memory of a very familiar place such as childhood/parent's home may be qualitatively different from recalling a list or video from a few weeks ago. This is especially likely to be the case given that participants were not told that they would be asked for such details. If the previous context cannot be mentally reinstated at all, then it would not produce a context change. This would lead to little or no forgetting.

In order to assess the possibility that participants were unable to access and reinstate the context of Session 1, I analyzed trends in the L0/Video recall data. My examination suggests that participants had great difficulty in accessing the Session 1 context. Specifically, the List group in both experiments started out by approximately equivalent recall rates in Session 1 (on the average, they recalled 72% of L0 during Session 1). However, In Experiment 1, participants that retrieved L0 between the two remaining lists in Session 2 recalled 55% of L0. In contrast, in Experiment 2, participants who retrieved L0 between the two remaining lists were able to retrieve only 23% of L0.

This is quite low recall of L0 despite it being cued via letter cues in Experiment 2 as opposed to free recall in Experiment 1. Thus, poor access to L0 episode in Experiment 2 could be one of the reasons for the observed null effects. Consistent with this argument, in Experiment 2, 11% of participants who retrieved L0 in Session 2 could not recall *any of the words* from L0, whereas in Experiment 1, all participants were able to recall at least one word from L0. Similar trends were observed also in the Video group. Specifically, in Experiment 1, all Video participants were able to retrieve at least some images and details from the video, whereas in Experiment 2, 23% of participants in the Video group either could not recall having watched a video *at all* or could not recall *any* of the details of the video. In other words, almost a quarter of participants in the Video condition in Experiment 2 forgot that they had seen a video, or at best, could not retrieve anything specific about that particular memory. It is interesting to note that twice as many Video participants (23%) recalled nothing about the video compared to participants in the List group, where 11% could not retrieve any L0 words. One possible explanation is that L0 was given multiple presentations along with an immediate test in Session 1, whereas the Video group was presented with video only once and was never tested on it except for answering three questions about it after watching it. Overall, if both the Video group and the List group experienced greater difficulty in reinstating Session 1 context, then they would have an equivalent L1 recall because they did not experience an internal context change.

Although the List group was more likely to retrieve Session 1 information than the Video group (many of who completely forgot that they had even seen a video), it is nevertheless possible that the List group was able to utilize the letter cues in successfully retrieving L0 items without necessarily successfully reinstating the Session 1 context. Thus, even though the List group's L0 recall rates are above floor, it is possible that they are not indicative of successful context change but are instead driven by the availability of letter-cues. Although previous research successfully used letter cues to help retrieve L0 (Sahakyan & Hendricks, 2012; Sahakyan & Smith, 2014), those studies did not incorporate any delay manipulation between the lists, and it is possible that a combination of delay and the provision of letter cues interact in some ways to circumvent context accessibility. If L0 is relatively recent, then the letter cues may play no active role in helping retrieve those items because L0 could be reinstated even without those letter cues. In contrast, when L0 is more "old", then participants are more likely to rely on the letter cues to help retrieve those items, and the more they rely on the letter cues, the less likely they may be to utilize contextual information (e.g., Raaijmakers & Shiffrin, 1980).

Overall, it is my conclusion that Experiment 1 was underpowered, which may have contributed to some of the null findings. It is well established that recall results are reliable indicators of context change and, ignoring the Video condition, Experiment 1 is a direct replication of a typical LBL experiment. Thus, the null recall results in Experiment 1 are the prime indication that more participants were needed in order to achieve statistical significance. Had more participants been included, it is possible that the time-estimation results could have also achieved significance—and, although this is only speculation, the obtained trend suggests a possible relationship. Given the typical effect sizes found in context change research, future research involving a similar experiment will require recruiting a relatively large sample of participants. An a priori power analysis for a repetition of the current research reveals that one should include 40 participants in each condition in order to obtain appropriate statistical power (disregarding outliers). Experiment 2 did not suffer from the same shortcoming. Instead, the results suggest that the context-change manipulations failed. Not only were there no group differences found in recall rates, none of the groups' time estimates differed from the objective duration. Future research should strive to include measures that would ensure that target context is more actively utilized and retrievable especially after a long delay.

REFERENCES

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval induced forgetting: evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7,* 522–530. doi: 10.3758/BF03214366

Block, R. A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4,* 530-544.

Block, R. A. (1992). Prospective and retrospective duration judgment: the role of information processing and memory. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, action and cognition: Towards bridging the gap* (pp. 141-152). Dordrecht, the Netherlands: Kluwer.

Block, R. A., & Reed, M. A. (1978). Remembered duration: evidence for a contextual -change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 656-665. doi:10.1037/0278-7393.4.6.656

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin and Review, 4,* 184-197.

Delaney, P. F., Sahakyan, L., Kelley, C. M., & Zimmerman, C. A. (2010). Remembering to forget: the amnesic effect of daydreaming. *Psychological Science, 21,* 1036-1042.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145-154. doi:10.1037/h0048509

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 112-127.

Jonker, T.R., Seli, P., & MacLeod, C. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review, 4,* 852-872.

Mensink, G. J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434-455. doi: 10.1037/0033-295X.95.4.434

Mulji, R., & Bodner, G. E. (2010). Wiping out memories: new support for a mental context change account of directed forgetting. *Memory, 18,* 763-773.

Pastotter, B., & Bauml, K. H. (2007). The crucial role of encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology, 5,* 977-982.

Poynter, D. W., (1983). Duration judgment and the segmentation of experience. *Memory and Cognition*, *11*, 77-82. doi:10.3758/BF03197664

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

Roediger, H.L., & Karpicke, J.D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science, 17,* 249-255.

Roy, M. M., & Christenfeld, N. J. S. (2007). Bias in memory predicts bias in future task duration. *Memory & Cognition, 35,* 557–564. doi:10.3758/BF03193294

Roy, M. M., & Christenfeld, N. J. S. (2008). Effect of task length on remembered and predicted duration. *Psychonomic Bulletin & Review, 15,* 202–207. doi:10.3758/PBR.15.1.202

Sahakyan, L., & Hendricks, C. (2012). Context change and retrieval difficulty in the list-before-last paradigm. *Memory and Cognition. 40,* 844–860. doi:10.3758/s13421-012-0198-0

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology, 28,* 1064-1072. doi:10.1037/0278-7393.28.6.1064

Sahakyan, L., & Smith, J. (2014). "A long time ago, in a context far, far away": Retrospective time estimates and internal context change. *Journal of Learning, Memory, and Cognition, 40,* 86-93.

Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science, 168,* 1601-1603. doi:10.1126/science.168.3939.1601

Smith, S. & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods, 42,* 292-301. doi:10.3758/BRM.42.1.292

Smith, S., & Vela, E., (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8,* 203-220.

Unsworth, N., Spillers, G. J., & Brewer, G. (2011). Dynamics of context-dependent recall: An examination of internal and external context change. *Journal Of Memory and Language, 66*, 1-16.

Unsworth, N., Spillers, G. J., & Brewer, G. (2012). Evidence for noisy contextual search: Examining the dynamics of the list-before-last recall. *Memory, 20*, 1-13. doi:10.1080/09658211.2011.626430

Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. In M. A. Pastor & 1.Artieda (Eds.), *Time. internal clocks and movement* (pp. 143 -164). Amsterdam: Elsevier, North-Holland,

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. *Memory and Cognition, 22,* 344-351. doi:10.3758/BF03200861

APPENDIX A

STIMULI USED IN EXPERIMENT 1 AND 2

LIST A	**LIST B**	**LIST C**
artist	beaver	baby
bench	bottle	candle
carpet	cattle	gazelle
flower	dress	guitar
knife	hotel	insect
monkey	leader	jacket
ostrich	market	mouse
person	quilt	opera
radio	street	paper
ticket	tiger	pupil
uncle	voter	snail
watch	zebra	sugar