Journal of Memory and Language 66 (2012) 1–16

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Dynamics of context-dependent recall: An examination of internal and external context change

Nash Unsworth [a,*], Gregory J. Spillers [a], Gene A. Brewer [b]

[a] *Department of Psychology, University of Oregon, United States*
[b] *Department of Psychology, Arizona State University, Tempe AZ, United States*

ARTICLE INFO

Article history:
Received 13 January 2011
revision received 6 May 2011
Available online 11 June 2011

Keywords:
Context-dependent memory
Recall dynamics

ABSTRACT

Retrieval dynamics in context-dependent recall were explored via manipulations of external and internal context in two experiments. Participants were tested in either the same or different context as the material was learned in and correct recalls, errors, and recall latency measures were examined. In both experiments changes in context resulted in fewer correct items being recalled than when context remained the same. However, the context change manipulation did not affect the number or type of recall errors or recall latency in either experiment. These results are consistent with the notion that changes in context result in a reduction in the associative strengths of items because there are fewer overlapping contextual features between encoded features and features present at test. Other potential mechanisms of context-dependent recall effects are discussed.

Introduction

> "An experience can be recalled most readily in those environmental situations with which it has the most direct, the strongest, and most numerous associations" (Carr, 1925, p. 251).
>
> "The learner is forming associations, not only intrinsic to the material which is being learned, but also between the parts of this material and the manifold features of the context or environment in which the learning is taking place. Two contexts must inevitably be present. One includes all of the stimulating conditions of the external environment; the other includes all intra-organic conditions. During time these contexts alter and it is at least highly probable that such alteration may remove the necessary eliciting stimulus" (McGeoch, 1932, pp. 365–366).

As the above quotes illustrate, the notion that context plays a role in remembering and forgetting has long been an important feature of many theories of memory. In particular, many theories of episodic memory assume that various contextual features or attributes are encoded along with content information during acquisition (e.g., Anderson & Bower, 1972; Estes, 1955; Gillund & Shiffrin, 1984; Hintzman, 1988; Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980; Tulving, 1983). These contextual features are then used to probe memory during retrieval in order to access the relevant information. When there is a strong overlap between the contextual features present at encoding and the contextual features present at retrieval, performance is high. When there is a mismatch between the contextual features at encoding and the features present at retrieval, performance is low. Despite much evidence for these basic findings, the underlying mechanisms by which this occurs are still not well understood. In particular, there are a number of potential ways in which context-dependent recall effects may arise. In the current study we examined context memory effects in free recall from a search model

* Corresponding author. Address: Department of Psychology, University of Oregon, Eugene, OR 97403, United States.
E-mail address: nashu@uoregon.edu (N. Unsworth).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2011.05.001

framework (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1980; Rohrer, 1996; Wixted & Rohrer, 1994) in order to better understand context-dependent memory.

Context-dependent recall

As noted above a number of studies have found that when the context at encoding matches that at retrieval performance tends to be high. When there is a mismatch, performance tends to be low. Of course what is meant by context can vary greatly depending on a number of factors. In the current study we will consider context to refer to the joint contribution of internal states and external features of the environment that are associated with the target items (Klein, Shiffrin, & Criss, 2007). These features can include environmental features like the room the experiment takes place in, the background of the computer monitor, the illumination conditions of the room, objects in the room the experiment takes place in, physical position in the room, temperature of the room, irrelevant auditory information in the room, as well as odors present in the room. These features can also include internal features such as the mood of the participant, the current cognitive state of the participant, changes in the participant's state due to mind wandering or lapses of attention (i.e., boredom), transitory changes in internal states such as brief feelings of hunger, as well as intentional shifts in strategies. These features in conjunction can be referred to as context and each of these are likely associated with the target information and can potentially influence the ability to accurately recall the target information. That is, changes in these various types of context can lead to context change effects.

Much of the early work on context dependent recall examined external changes in context. For example, work by Abernethy (1940) suggested that changes in the room where information was studied versus the room where testing occurred can lead to impairments in recall for students (see also Weir & May, 1988; but see Saufley, Otaka, & Bavaresco, 1985). Follow-up research has generally corroborated these basic results suggesting that changes in external context leads to impairments in recall. Changes in external context including changes in room (e.g., Smith, Glenberg, & Bjork, 1978), physical position (Rand & Wapner, 1967), physical environment (Godden & Baddeley, 1975), presence or absence of odors (e.g., Schab, 1990), presence or absence of music (e.g., Smith, 1985), changes in experimenter (e.g., Smith & Vela, 2001) as well as changes in the color background of the monitor (e.g., Isarida & Isarida, 2007), all have been found to lead to context change effects under some circumstances (see Smith & Vela, 2001 for a review).

In addition to demonstrating context change effects with changes in external context, other studies have demonstrated context change effects with changes in internal context. Specifically, several studies have suggested that changes in mood states can lead to reductions in recall under some circumstances (e.g., Bower, 1981; Bower, Monteiro, & Gilligan, 1978). Similarly, research examining the role of psychopharmacological states on recall has found that when the learning and test states match, performance is typically better than when they mismatch. This effect has been found in recall when participants ingest alcohol and when participants ingest marijuana (see Eich, 1980 for a review). In both cases, the psychopharmacological state at test acts as a context cue that is used to probe memory for the list of items suggesting that state-dependent recall is a result of the cue-dependent nature of memory (Bower, 1981; Eich, 1980; Tulving, 1983).

Recent work has begun to examine how other internal contextual features can influence recall performance. For instance, Sahakyan and Kelley (2002) recently proposed a context change account of list-method directed forgetting. In list-method directed forgetting, participants learn a list of words and then are either told to remember or forget those words. They are then presented with a second list of words to remember. At test participants are told to recall both lists of words. Directed forgetting is found when participants who were told to forget the first list have lower recall performance for first list items than participants who were told to remember the first list. According to Sahakyan and Kelley, directed forgetting occurs because once participants are told to forget the first list, those participants actively change their internal context by thinking about something else. At recall, the current context is a weak cue and does not sufficiently activate the first list items leading to forgetting. To examine this notion, Sahakyan and Kelley had participants perform a traditional directed forgetting task with one group told to remember the first list and one group told to forget the first list. In addition, Sahakyan and Kelley had a context change condition in which after the presentation of the first list of items, participants were instructed to imagine being invisible and to write (in 45 s) the things they would do if they did not have to take responsibility for their actions. Consistent with their context change account, the authors found reduced recall for first list items when their internal context was changed. In a subsequent study, the authors found similar results when participants were instructed to imagine their parents' house and to draw a layout of their parents' house. Again performance was reduced compared to the control group who did not change their internal context by thinking about their parents' house. However, if participants were instructed to mentally reinstate their context, performance increased. Thus, internal context can be changed via simply thinking about things unrelated to the current experimental situation and this internal context change can result in decreased recall performance (see also Delaney, Sahakyan, Kelley, & Zimmerman, 2010; Pastötter & Bäuml, 2007).

Collectively prior research suggests that changing either external or internal context features can lead to reductions in recall. Theoretically, changes in either external or internal contextual features result in similar decrements due to the fact that both types of context are associated with item/content information at encoding and thus both are useful in terms of cuing/probing memory at test. Despite these apparent similarities, little work has been done to directly examine whether external and internal context change leads to similar decrements in recall and similar patterns of output. That is, not only is there scarce empirical work examining similarities and differences between external and internal context change effects in recall, but there is also little work examining recall performance as a function of dependent measures other than mean proportion correct. Although mean levels of accuracy provide important information about what is recalled, more fine-grained analyses are needed to better examine the dynamics of context-dependent recall and to assess different potential theoretical mechanisms for context-dependent recall.

Search processes and the dynamics of free recall

The work reviewed thus far has focused almost exclusively on probability of correct recall. However, an examination of recall latency can also be informative in terms of better understanding how participants search for target items in free recall tasks. Recall latency refers to the time point during the recall period when any given item is recalled, and mean recall latency is simply the average time it takes to recall items. For instance, if items are recalled 5 s, 10 s, and 15 s into the recall period, mean recall latency would be 10 s. Prior work has suggested that recall latency distributions provide important information on the dynamics of free recall. In particular, prior (Bousfield & Sedgewick, 1944; Indow & Togano, 1970; McGill, 1963; Roediger, Stellon, & Tulving, 1977; Rohrer & Wixted, 1994; Wixted & Rohrer, 1994) work has suggested that cumulative recall curves are well described by a cumulative exponential

$$F(t) = N(1 - e^{-\lambda t}),$$

where $F(t)$ represents the cumulative number of items recalled by time t, N represents asymptotic recall, and λ represents the rate of approach to asymptote. Thus, if given enough time to recall, N should equal (or be roughly equal to) the number of items recalled (or probability of recall). However, these items can be recalled either quickly or slowly and this information is captured by λ. Specifically, when items are recalled quickly during the recall period λ is relatively large, whereas when items are recalled slowly during the recall period λ is relatively small. Thus, cumulative recall curves provide information not only on how many items are recalled, but also information on how quickly those items are retrieved. Importantly, overall mean recall latency is simply the inverse of λ when the cumulative functions are perfectly exponential (e.g., Wixted, Ghadisha, & Vera, 1997), and thus it is possible to either estimate recall latency from λ or to compute it directly from the latencies associated with each recalled item.

Overall recall latency distributions are consistent with search models of free recall (Rohrer, 1996; Shiffrin, 1970). In these models it is assumed that during recall a retrieval cue activates a subset of representations in memory that are related to the cue in some fashion. This delimited subset is known as the *search set* and during recall, item representations are sampled (with replacement) from the search set based on a relative strength rule (Raaijmakers & Shiffrin, 1980; Rohrer, 1996; Shiffrin, 1970). Specifically, in search models of this type the probability of sampling any particular item is equal to the strength of the item divided by the sum of all item strengths within the search set (e.g., $s_i/\Sigma s_j$). After an item has been sampled, it must then be recovered into consciousness. In these search models, recovery of an item depends on the item's absolute strength rather than on its relative strength. Specifically, items whose strength exceeds some critical threshold will be recovered and can be recalled, whereas weak items that do not exceed the threshold will not be recovered (Rohrer, 1996). Important for models of this type is the notion that all items can be sampled, but only those items whose strength exceeds the threshold can actually be recalled. Thus, it is possible to differentiate these two aspects of recall (sampling and recovery). Finally, after an item has been recovered, it is subjected to a monitoring and editing process that determines whether the item is correct and recalled, or incorrect and not recalled.

According to search models of this type, N reflects the number of target items in the search set whose absolute strength exceeds some threshold (i.e., the numerator in the relative strength rule; e.g., Rohrer, 1996). Recall latency, and λ, reflects the number of items within the search and thus reflects relative strength (i.e., the denominator in the relative strength rule). The larger the search set the longer on average it will take to recall any given item. Important, evidence for this type of model as well as for a distinction between N and λ comes from a number of studies that have manipulated aspects of free recall and found that some variables affect N, but have no effect on λ, whereas other variables seem to primarily affect λ. For instance, Rohrer and Wixted (1994) manipulated presentation duration and found that this manipulation affected the number of items recalled (N), but had no effect on recall latency (λ). Consistent with search model explanations of the presentation duration (e.g., Gillund & Shiffrin, 1984) this is because presentation duration influenced the absolute strength of each item, but did not affect the relative strength of the items (i.e., all items had the same boost in strength and thus relative strength was unchanged; see also Wixted et al. (1997)). In another experiment Rohrer and Wixted (1994) manipulated list-length and found that as list-length increased, the number of items recalled increased (although probability correct decreased), and recall latency increased (see also Unsworth, 2007). This is consistent with the notion that as list-length increased, relative strength decreased leading to a drop in probability of recall and an overall increase in recall latency. Further evidence consistent with this notion comes from a study by Wixted and Rohrer (1993) that examined the build and release of proactive interference. In this study, Wixted and Rohrer found that as proactive interference increased and probability of recall subsequently decreased, overall recall latency increased. Similar to the list-length effects, this is presumably because as proactive interference built up, more items were included in the search set and relative strength decreased (i.e., the denominator increased in the relative strength rule). Thus, although N decreased this was due to a change in relative strength rather than absolute strength given that the search set was likely composed of both target items and intrusions from prior lists. Indeed, in a recent large-scale individual differences study

Unsworth (2009) found that recall latency and number of intrusions were positively correlated, and both were negatively correlated with recall accuracy. This suggests that the inclusion of intrusion errors into the search set causes an overall increase in search set size leading to a lower probability of sampling target items and increase in the average time to sample target items. Collectively, the results from these studies suggest that recall latency provides an index of overall search set size (see also Shiffrin, 1970).

It should be noted that this simple random search model assumes that items are randomly sampled and recalled. Clearly there are non-random forces at play resulting in serial position functions, probability of first recall functions, and lag-recency effects (see below). However, the random search model has been validated by prior research suggesting that despite non-random recall the overall interpretation provided by the random search model still holds (e.g., Rohrer, 1996; Vorberg & Ulrich, 1987; Wixted & Rohrer, 1994). Thus, the random search model is a useful tool for interpreting the effect of various manipulations on recall latency. More complex search models that allow for variable items strengths, inter-item associations, and strategic search processes like search termination rules would likely make similar predictions as the simple random search model, but these models could also provide slightly different interpretations (e.g., Raaijmakers & Shiffrin, 1980). Important for the current study, examining the dynamics of free recall (using the random search model) can provide valuable information regarding how certain manipulations, like context change, influences recall performance by impacting different theoretical mechanisms like sampling and recovery.

The present study

The goals of the present investigation were threefold. First, we wanted to examine multiple indicators of performance in context-dependent recall beyond differences in mean recall. Specifically, we wanted to break down context-dependent memory effects in recall by examining overall serial position effects, probability of first recall effects, as well as conditional response probabilities. As noted previously, most prior work examining context-dependent recall effects has usually only examined mean levels of performance. However, it is well known that various manipulations can be localized to certain serial positions. Thus, we wanted to see if context-dependent recall effects could be localized to certain serial positions or whether these effects occur across all serial position effects. Some initial work by Nixon and Kanak (1981) suggests that context-dependent recall effects partially arise from differences in the recency portion of the serial position curve. Similar results were recently reported by Isarida and Isarida (2006) using a version of the continuous distractor task in which the target items were pairs of words. Thus, initial work seems to suggest that context-dependent recall effects might be localized to recency portions of the serial position curve. Although examinations of overall serial position functions can be informative, these functions can be further broken down into examinations of probability of first recall and conditional response probabilities (e.g., Kahana, Howard, Zaromb, & Wingfield, 2002). Probability of first recall (PFR) refers to the number of times the first word recalled comes from a given serial position divided by the number of times the first recalled word could have come from that serial position. For instance, if a person begins recall with the last presented word nine out of ten times, then the probability of first recall for that serial position would be .90. PFR provides important information about how participants initiate recall at test through the use of context cues. If context-dependent recall effects arise from differences in how context cues items initially we should see differences in the PFR functions. Once an item is recalled, typically the next item recalled comes from a nearby serial position and this effect occurs more strongly in the forward diction than the backward direction (Howard & Kahana, 1999; Kahana, 1996). This effect is known as the lag-recency effect and it can be measured via an examination of condition response probabilities (CRPs). To date, only one study has examined lag-CRP effects for context-dependent memory (Isarida & Isarida, 2006). In this study, Isarida and Isarida found nearly identical lag-CRP functions for items recalled in the same and different contexts suggesting that context-dependent recall did not influence lag-CRP effects. However, this study specifically examined paired recall of items presented as pairs, and it is not known what these effects will look like in a more traditional free recall task.

In addition to examining serial position functions, PFR functions, and lag-CRP functions for context-dependent recall, we also wanted to examine recall latency and inter-response times (IRTs) for context-dependent recall. As noted previously, an examination of cumulative recall functions and mean recall latency (as well as mean IRTs) can provide valuable information about the dynamics of context-dependent recall and the possible underlying mechanism. Thus, our second goal in the current investigation was to utilize information gleaned from recall latency to attempt to distinguish potential mechanisms that underlie context-dependent recall effects. In particular, there are several theoretical possibilities for context-dependent recall effects. First, it is possible that changes in context lead to a reduction in the associative strengths of items because there are fewer overlapping contextual features between encoded features and features present at test. According to search models like the search of associative memory (SAM) model this reduction in feature overlap results in lower recovery probabilities making it less likely that sampled items in the context change condition will actually be recovered (Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980; see also Spear & Riccio, 1994). Thus, changes in context are similar to changes in presentation duration in that both seem to affect recovery probabilities, but do not influence sampling probabilities (Gillund & Shiffrin, 1984). In terms of cumulative recall functions this would mean that the difference between conditions in which context is the same or different should result in differences in N, but no difference in λ and recall latency. According to this account, this would occur because the same number of items are included in each search set, but what differs is the fidelity of the items in the search set. Thus, participants should be searching through the same number of items leading to equivalent recall latencies, but participants in the context change condition will have fewer intact items and will therefore have lower overall levels of recall. Overall, then this account predicts that context change effects should be similar to manipulations of presentation duration in that differences should arise in the total number of items recalled but no differences should occur for recall latency (Rohrer & Wixted, 1994).

A second possible reason for context change effects is that context change may lead to an increase in the number of items within the search set due to the fact that participants might be exposed to intervening material (e.g., other items, thoughts, events, or stimuli) before being tested (Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1988; Spear & Riccio, 1994). For example, Spear and Riccio (1994) suggested that "new stimuli noticed upon a change in context may arouse competing memories that interfere with retention of the target memory" (p. 57). This reason is particularly important when other items have been presented after the initial items and/or when the delay between presentation and test increases differentially across conditions. Indeed, Pastötter and Bäuml (2007) have recently argued that context-dependent recall effects occur only when interfering material is presented after the target list. Thus, context-dependent recall occurs because a change in context results in other information being included in the search set. To model this notion, Gillund and Shiffrin (1984) assumed that when context changes (based on an increase in delay) additional irrelevant items (i.e., junk memoires) are included in the search set. The net effect of this is the same as increasing list-length or proactive interference in that sampling probabilities decrease, but recovery probabilities are not affected (Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1988; Wixted & Rohrer, 1993). In terms of cumulative recall functions this would mean that difference between conditions in which context is the same or different should result in differences in N, and difference in λ and recall latency such that recall latency is longer in the context change condition than when context remains the same. According to this account, this would occur because as the overall search set size increases the probability of finding a correct item amongst irrelevant items decreases. Thus, not only is it less likely that a correct item will be sampled, but this also means that it should take longer on average to find correct items than when the search set is smaller (e.g., Rohrer & Wixted, 1994; Unsworth, 2007). Furthermore, if irrelevant information is being included in the search set, we should also find that participants in the context change condition recall more intrusions than participants in the same context. Overall, then this account predicts that context change effects should be similar to manipulations of list-length and proactive interference in that differences should arise in the total number of items recalled, in recall latency, and in the number of intrusions emitted (Gillund & Shiffrin, 1984; Rohrer & Wixted, 1994; Unsworth, 2007).

In addition to these two possibilities, there are other possible ways in which changes in context might influence recall. For instance, Smith (1994) suggested that changes in context might influence the size of the search set such that with changes in context fewer target items are actually included in the search set compared to when context is not changed. That is, with a large enough change in context, items with weak contextual associations with the current context will not be included in the search set resulting in a smaller than normal search set. Thus, the reduction in levels of recall occurs because target items in context change conditions are not included in the search set and thus, are not even sampled. In terms of cumulative recall functions this would mean that differences between conditions in which context is the same or different should result in differences in N, and difference in λ and recall latency such that recall latency is shorter in the context change condition than when context remains the same. This is because according to this account the overall search set size is reduced and thus, it should take less time on average to find those target items that are actually included in the search set.

These three search model explanations of context dependent recall all predict that N should be lower when context is changed than when it is the same, but these three explanations make different predictions about λ when context is changed. To get a better sense of these predictions, we generated cumulative recall functions for a context same condition and compared this to the three context change possibilities. Shown in Fig. 1 are the resulting curves. It can be seen that the three context change possibilities have similar asymptotic levels of recall, which are lower than the context same condition, but the context change possibilities differ in the rate of approach to asymptote. For the context same condition the resulting parameter estimates were $N = 16$ and $\lambda = .12$. For the context change possibility in which there are fewer recoverable targets (labeled Nonrec in the figure) the resulting parameter estimates were $N = 12$ and $\lambda = .12$. Thus, changes in context resulted in fewer items being recalled, but these items were recalled at the same rate as those in the context same condition. For the context change condition in which there are irrelevant items included in the search set (labeled Large in the figure) the resulting parameter items were $N = 12$ and $\lambda = .08$, suggesting that when more items are included in the search set it should take longer to reach asymptotic levels. Finally, for the context change possibility in which there are fewer items included in the search set (labeled Small in the figure) the resulting parameter estimates were $N = 12$ and $\lambda = .16$, suggesting that when there are fewer targets included in the search set asymptotic levels will be reached quickly. Thus, the three different possibilities make similar predictions in terms of overall number of items recalled, but they differ in terms of the rate at which those items will be recalled suggesting that an examination of cumulative recall curves and recall latency is needed to better examine these possibilities.

Another possible reason for context change effects (and one not mutually exclusive to the prior possibilities) is that perhaps changes in context change the way participants search memory and use previously recalled items to cue new items. That is, in many search models of free recall it is assumed that once an item is recalled that item plus

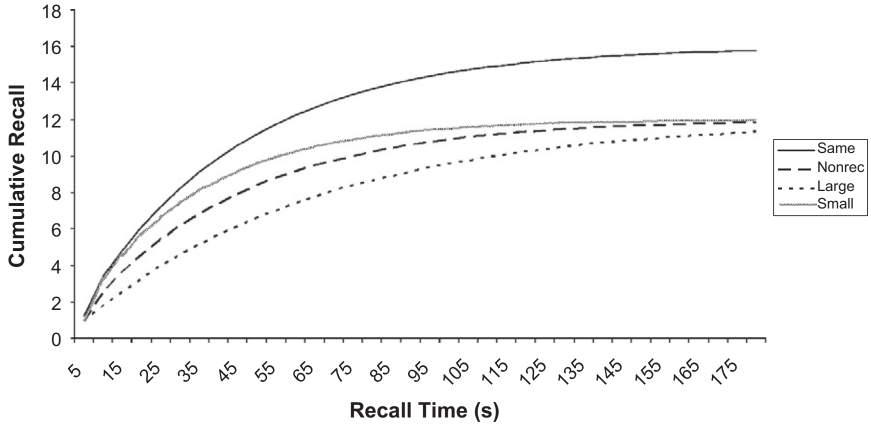

Fig. 1. Predicted cumulative recall functions for a context same condition and for three context change conditions. See text for details.

overall context are used to cue the next item (e.g., Howard & Kahana, 2002; Raaijmakers & Shiffrin, 1980). Lag-CRP effects provide important evidence for this process. According to context retrieval models like the Temporal Context Model (TCM) these effects are indicative of a contextual retrieval process where items are sampled based on the context present at test as well as context associated with the just recalled items. Thus, once item 5 is recalled its context is used to search and retrieve subsequent items. Items that share contextual elements with the just recalled item (i.e., item 6) will then have a higher probability of being sampled than items that share few contextual elements with the just recalled item (i.e., item 10). In terms of context change effects it is possible that reductions in overall levels of recall are due, in part, to a weakening of associative contextual bindings between items such that once an item is recalled it acts as a poor cue for subsequent items. Furthermore, it may be the case that changes in context lead to changes in participants retrieval plans such that participants used the encoding context to link items together and when context changed that strategy was disrupted and could not be utilized. In both cases, only overall context would serve as a cue (and a poor one at that) resulting in not only lower levels of recall but drastically reduced lag-CRP effects. That is, if participants are only using overall context to cue their memory system then the lag-CRP functions should be flat or dramatically reduced suggesting that overall search was fairly random in nature. If the search is fairly random in nature, one would also expect flat serial position functions and flat PFR functions. Thus, rather than examining cumulative recall functions this possibility will be examined via an examination of the other indicators of performance.

The final goal of the present study was to examine potential similarities and differences between external and internal context change. As noted previously, both external and internal context change have been found to influence recall performance. Generally it is assumed that these two forms of context change arise from similar theoretical mechanisms. However, little work has directly compared external and internal context change manipulations and no prior work has examined these two forms of context change in terms of other indicators of performance as specified in the current study. If external and internal context change represent fundamentally the same cognitive operations one would expect for them to demonstrate the same pattern of results across all dependent measures. If there are fundamental differences between these two forms of context then differences should appear and why they appear should be informative.

To examine these issues we conducted two experiments in which participants learned a list of 40 words and then recalled those words in either the same or different context. In Experiment 1 we examined external context via a change in room manipulation. In Experiment 2 we examined internal context by having some participants think about their parents' house while others engaged in a distracting task (Sahakyan & Kelley, 2002). In both experiments we not only examined overall recall levels but also examined serial position functions, PFR functions, lag-CRP functions, cumulative recall functions, recall latency, as well as IRTs. A more fine-grained analysis of context change effects should allow for a better understanding of how changes in context influence memory and the possible mechanisms that underlie these effects.

Experiment 1

The purpose of Experiment 1 was to examine external context change in free recall. Participants learned a list of 40 words in one room and were either tested in that same room or tested in a different room. Performance was examined based on a number of dependent measures. As noted previously, if external context change effects are due to differences in recovery we should see that overall recall levels differ between the two conditions, but there are no differences in recall latency. If external context change effects are due to differences in the size of the search set we should see that overall recall levels differ and recall latency is either longer in the context change condition (due to the inclusion of irrelevant or junk items), or is shorter in the context change condition (due to fewer target items actually being included in the search set). Finally, if external context change effects are due to changes in contextual associations between items we should see that overall recall is reduced in the context change condition and lag-CRPs, serial position functions, and PFR functions are relatively flat (or drastically reduced).

Method

Participants and design

Participants were 38 (19 in each condition) undergraduate students recruited from the subject-pool at the University of Georgia. Participants were between the ages of 18 and 35 and received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately thirty minutes. Participants were randomly assigned to either the context same or context change condition. Participants learned a list of 40 nouns in one room and then were either tested in the same or different room. Words were 40 nouns selected from the Toronto word pool (Friendly, Franklin, Hoffman, & Rubin, 1982).

Procedure

Participants were told that they would be presented with a series of words and that their task was to try to remember the words for a later test. Before beginning, participants completed a practice list to familiarize themselves with the task. The practice list consisted of a series of 10 letters presented at 2 s. At recall, participants saw three question marks (???) appear in the middle of the screen indicating that the recall period had begun. Participants had 60 s to recall as many of the letters from the practice list as possible in any order they wished. Participants typed their responses and pressed Enter after each response clearing the screen. Prior to the practice and real trials, participants received a brief typing exercise (typing the words one-ten) to assess their typing efficiency. For the real recall task, participants learned 40 words with each word being presented alone for 2 s. All participants learned the list of words in a brightly illuminated room on the fifth floor of the Psychology Department in which various objects were placed. These objects included: a fake deer head, a lucky cat statue, a parrot kite, as well as a small statue of David. After learning the list of words all participants were then taken down to the third floor of the Psychology Department. Those participants in the context change condition were then lead to a dark room on the third floor to recall the words. Included in this room were a computer, a desk, and a chair. Participants in the context same condition were taken back up to the fifth floor to recall the words in the same room as the words were learned. The reason for taking all participants to the third floor was to equate for possible differences in psychological disruption which has been proposed as a possible alternative explanation for context change effects (e.g., Nixon & Kanak, 1981; Strand, 1970). Thus, all participants were exposed to the same disruption. The inter-task interval was 2 min for all participants. Once placed into the appropriate room, all participants were instructed to recall as many words as possible from the list in any order they wished. Participants typed their responses and were given 3 min for recall.

Results

Correct recalls

Consistent with prior research there was a robust context change effect in that participants who recalled the words in the same context as they were learned recalled more words ($M = .31$, $SE = .02$) than participants who recalled in a different context ($M = .23$, $SE = .02$), $t(36) = 2.28$, $p < .05$, $\eta^2 = .13$. As shown in Fig. 2a, examining proportion correct as a function of serial position suggested a main effect of serial position, $F(9, 324) = 6.94$, $MSE = .06$, $p < .01$, partial $\eta^2 = .16$, such that there was a strong primacy effect but no recency effect consistent with other work using delayed free recall (Glanzer & Cunitz, 1966). Note for analyses we combined every four serial positions. In terms of differences between the context groups, as shown in Fig. 1a, the groups demonstrated equivalent serial positions effects and there was no context group by serial position interaction, $F(9, 324) = .51$, $MSE = .06$, $p > .86$, partial $\eta^2 = .01$. Thus, although there was a main effect of context, the two groups demonstrated largely similar serial position functions.

Moving beyond serial position effects we next examined how participants initiated recall and transitioned to new items during the recall period by examining PFR and lag-CRP functions. Shown in Fig. 2b are the PFR functions for both conditions. As can be seen there was a main effect of serial position, $F(9, 324) = 4.54$, $MSE = .09$, $p < .01$, partial $\eta^2 = .11$, suggesting that participants primarily initiated recall with primacy items. Furthermore, as shown in Fig. 2b, the PFR functions did not differ for context change and context same conditions, $F(9, 324) = .51$, $MSE = .09$, $p > .87$, partial $\eta^2 = .01$, suggesting that participants in both conditions initiated recall in a similar fashion. Shown in Fig. 2c are the lag-CRP functions for forward and backward transitions. These functions represent the conditional response probability (CRP) of forward and backward transitions made between correctly recalled items based on the presentation lag. These CRP functions were calculated exactly the same way as previous research has done (Howard & Kahana, 1999; Kahana, 1996). Consistent with this previous research (Howard & Kahana, 1999; Kahana, 1996) the majority of transitions were of a lag of 1 and in the forward direction. Specifically, transitions associated with a short lag were more likely than transitions associated with a long lag, $F(4, 144) = 10.30$, $MSE = .004$, $p < .01$, partial $\eta^2 = .22$, and the lag effect was stronger in the forward than backward direction, $F(4, 144) = 2.96$, $MSE = .005$, $p < .01$, partial $\eta^2 = .08$. Thus, participants tended to begin recall with the first word presented in a list and then tended to recall items in the forward direction leading to large primacy and virtually no recency effects. As shown in Fig. 2c, there were no differences between the context conditions in terms of the lag-CRP functions. Specifically, neither the lag by condition interaction nor the direction by lag by condition

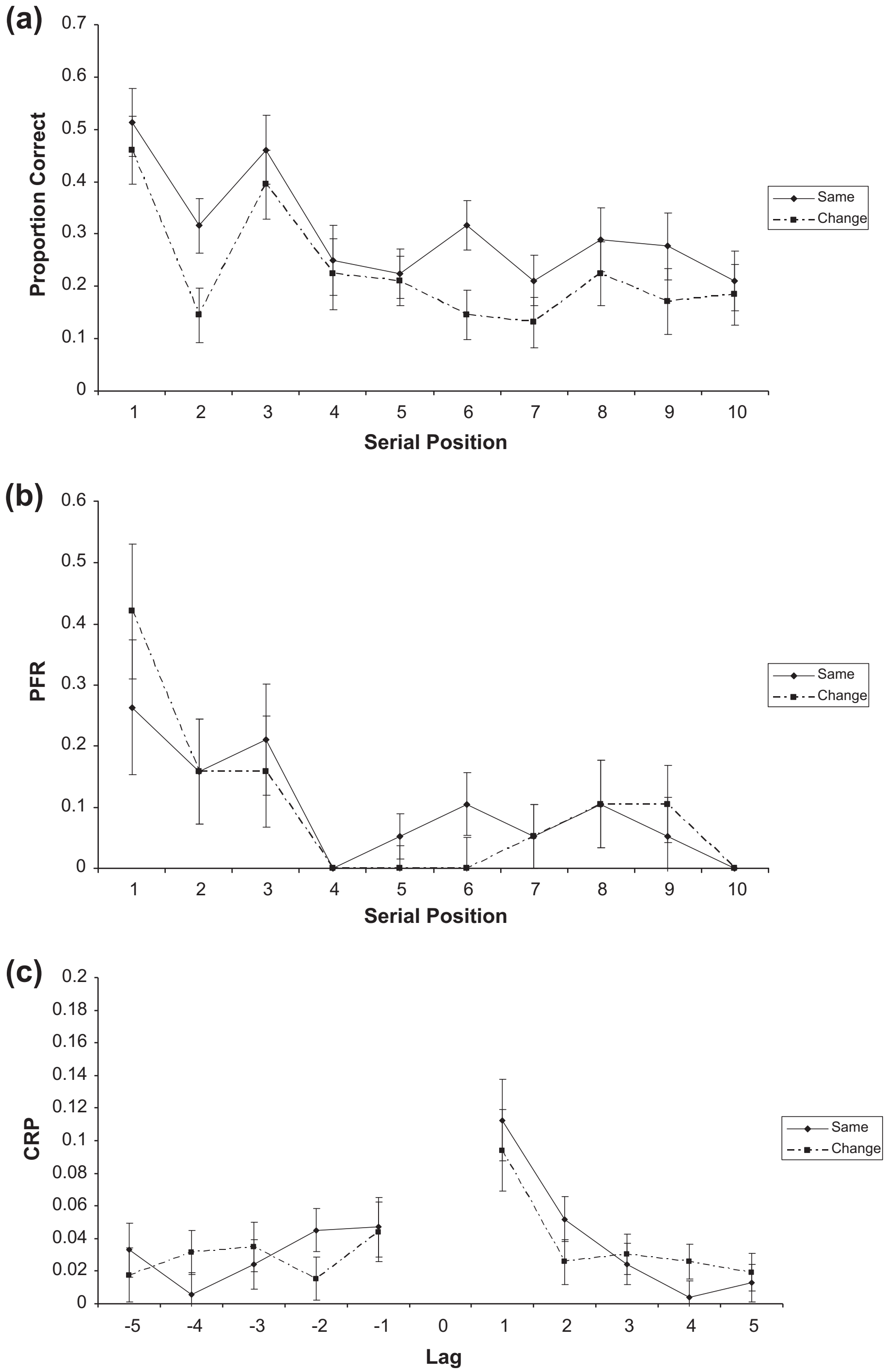

Fig. 2. (a) Proportion of correct recall as a function of serial position and context condition. (b) Probability of first recall (PFR) as a function of serial position and context condition. (c) Conditional response probability functions for forward and backward transitions per list as a function of lag and context condition. Error bars represent one standard error of the mean.

interaction reached conventional levels of significance, all *p*'s > .13. Thus, despite the fact that participants in the context same condition recalled more items than participants in the context change condition, these groups demonstrated remarkable similarities in terms of how they initiated recall and transitioned between correct recalls.

Error responses

Next, error responses were examined to better understand the recall process. Errors were classified as intrusions (items not presented on the list) or repetitions (items from the list that had already been recalled). Overall the two conditions made similar numbers of errors. Specifically, although participants in the context same condition had numerically more intrusions ($M = 3.21$, $SE = .69$) than participants in the context change condition ($M = 1.79$, $SE = .55$) this difference did not reach conventional levels of significance, $t(36) = 1.59$, $p > .12$, $\eta^2 = .06$. Additionally, there was no difference in repetitions between the context same ($M = .58$, $SE = .25$) and the context change conditions ($M = .58$, $SE = .18$), $t(36) = .00$, $p > .99$, $\eta^2 = .00$. Thus, the changes in context did not influence the number or type of errors generated.

Latency measures

In addition to the above measures on correct recalls and error responses, latency and IRT information was also examined to better understand the dynamics of delayed free recall. Previous research has shown that an examination of overall recall latency as well as inter-response times (IRTs) provides a window into the search process (Rohrer & Wixted, 1994; Unsworth, 2007; Wixted & Rohrer, 1994). Shown in Fig. 3 are the cumulative recall curves for the context same and context change conditions. This curve represents the cumulative number of items recalled for every 5 s during recall period and provides an overall depiction of the full time course of recall during the recall period. Consistent with previous research (Rohrer & Wixted, 1994; Wixted & Rohrer, 1994), the cumulative recall curve was well described by a cumulative exponential. As shown in Fig. 3 the symbols represent the data and the lines represent the best fitting cumulative exponential. The resulting parameter estimates suggest that the context same and context change conditions differed in N, but did not seem to differ in λ. Specifically, for the context same condition the parameter estimates were $N = 15.95$ and $\lambda = .10$. For the context change condition the parameter estimates were $N = 11.41$ and $\lambda = .11$. Additionally, the fits were acceptable with the functions each accounting for 99% of the variance and Kolmogorov–Smirnov tests were non-significant ($p > .61$).[1] Thus, the cumulative recall functions suggest that the two conditions differ in the total number of items recalled but the rate of approach to asymptotic levels did not differ.

Although the cumulative recall curve provides a general depiction of recall latency, a more detailed analysis of recall latency and IRTs is necessary to more fully understand recall dynamics. Therefore, recall latency and IRTs for correct recalls were examined in more detail. As shown in Table 1, there were virtually no differences between the two context conditions in terms of the recall latency measures. Specifically, in both conditions the first item was emitted around 5 s after the onset of the recall signal consistent with the notion of a pause preceding output, and there was no difference between the conditions, $t(36) = .63$, $p > .53$, $\eta^2 = .01$. Overall, average recall latency was between 35 and 40 s suggesting that, on average, participants emitted their responses 35–40 s into the recall period. This did not differ significantly between conditions, $t(36) = .96$, $p > .34$, $\eta^2 = .02$. Finally, overall IRTs were approximately 9 s suggesting that on average the time between recalls was 9 s. This also did not differ significantly as a function of condition, $t(36) = .76$, $p > .45$, $\eta^2 = .02$. Collectively, these results suggest that although there were differences in the total number of items recalled in the two conditions, there were no differences in the dynamics of recall as measured by various recall latency measures. Thus, the time to output responses was the same in the two conditions.

Discussion

The results from Experiment 1 were relatively straightforward. A change in external context resulted in fewer items being recalled than when context remained the same. Despite differences in overall levels of recall the change in context did not change the serial position function, the PFR function, or the lag-CRP function suggesting that participants regardless of condition largely recalled primacy items, recalled primacy items first, and transitioned in a forward manner to temporally close items. Furthermore, despite differences in overall recall levels, the conditions did not differ on any of the recall latency measures. Thus, the way in which participants recalled items and the rate at which they recalled items was the same despite the change in context. This overall pattern of results is most consistent with the notion that changes in context affect the recoverability of items. That is, if changes in external context result in fewer overlapping contextual features between context at encoding and context at test, then overall recovery levels should be lower. Such reductions in recovery would result in fewer items being recalled, but would not result in any changes in how those items were recalled. Furthermore, if changes in external context influence recoverability without influencing sampling probabilities, one would expect that although differences would arise in the number of items recalled, there would be no differences in the rate at which items are recalled (Gillund & Shiffrin, 1984; Rohrer & Wixted, 1994). Thus, the current results are inconsistent with the notion that changes in context affect sampling by either increasing or decreasing the size of the search set (i.e., the sampling space). If changes occurred in sampling, one would expect recall latency to be affected. However, there were virtually no differences between the conditions on any of the recall latency measures. Finally, if changes in context somehow led to changes in recall initiation or to reductions in associative bindings between items, one would expect differences in the PFR functions and differences in the lag-CRP functions. However, the two conditions produced nearly identical PFR and lag-CRP functions. Thus, the fact that differences were found in the number of items recalled, but not in how the items were recalled or in the rate of recall, the current results seem most consistent with the notion that changes in external context result in changes in

[1] Note, that although the cumulative exponential fit the data well, there are clear systematic deviations of fit. Specifically, the cumulative exponential initially underestimates and then over estimates the data. As shown by Vorberg and Ulrich (1987) this pattern is expected when item strengths vary. Given there is evidence in the current data for varying item strengths this pattern is to be expected. Despite these variations the simple search model still provides a useful interpretation of the data.

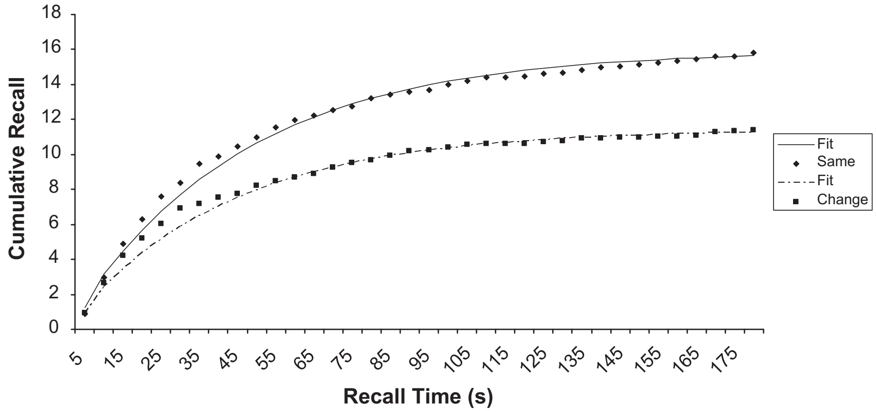

Fig. 3. Cumulative recall curves as a function of recall time and context condition. Symbols represent the observed data and the solid line represents the best fitting exponential.

Table 1
Recall latency measures (in s) as a function of context condition in Experiment 1.

Condition	Measure		
	Time-to-first	Recall latency	IRT
Same	5.09 (.61)	40.68 (4.21)	8.74 (.84)
Change	4.59 (.51)	35.49 (3.37)	9.59 (.77)

Note: IRT = inter-response time. Standard errors are shown in parentheses.

the recoverability of items. This notion is further examined in Experiment 2.

Experiment 2

The purpose of Experiment 2 was to replicate and extend the findings from Experiment 1 by examining internal context change in free recall. Participants learned a list of 40 words and then either changed their internal context by thinking and recalling the details of their parents' house (Sahakyan & Kelley, 2002) or maintained their internal context by simply performing an unrelated distractor task. Prior research has suggested that this type of diversionary task results in context-dependent recall effects due to an internal change in context (e.g., Delaney et al., 2010; Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002). Although this type of internal context change is somewhat different from changes of mood or psychopharmacological state, prior work has suggested that these diversionary tasks should lead to similar changes in internal context. This line of reasoning was explored in Experiment 2. As with Experiment 1 performance was examined based on a number of dependent measures. The same hypotheses and possibilities explored in Experiment 1 were examined in Experiment 2.

Method

Participants and design

Participants were 48 (24 in each condition) undergraduate students recruited from the subject-pool at the University of Georgia. Participants were between the ages of 18 and 35 and received course credit for their participation. Each participant was tested individually in a laboratory session lasting approximately thirty minutes. Participants were randomly assigned to either the context same or context change condition. Participants learned a list of 40 nouns in one room and then were either tested in the same or different room. Words were 40 nouns selected from the Toronto word pool (Friendly et al., 1982).

Procedure

Participants were told that they would be presented with a series of words and that their task was to try to remember the words for a later test. Before beginning, participants completed a practice list to familiarize themselves with the task. The practice list consisted of a series of 10 letters presented at 2 s. At recall, participants saw three question marks (???) appear in the middle of the screen indicating that the recall period had begun. Participants had 60 s to recall as many of the letters from the practice list as possible in any order they wished. Participants typed their responses and pressed Enter after each response clearing the screen. Prior to the practice and real trials, participants received a brief typing exercise (typing the words one-ten) to assess their typing efficiency. For the real recall task, participants learned 40 words with each word being presented alone for 2 s. After learning the list of words half of the participants engaged in a context change task while the other half performed a distractor task (Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002). Those participants in the context change condition were instructed to imagine their parents' house and to mentally walk through it. These participants were instructed to describe and draw the layout of their parents' house on paper for 46 s. The instructions for this context change manipulation were exactly the same as those used by Sahakyan and Kelley (2002) who have suggested that such a manipulation provides an effective means of changing internal context. Participants in the context same condition engaged in a 46 s distractor task before recall: Participants saw 23 three-digit numbers appear for 2 s each, and were required to write the digits in ascending order (e.g., Rohrer & Wixted, 1994; Unsworth, 2008). Following the context change task or the distractor task, all participants were instructed to recall as many words as possible from the list in any order they wished. Participants typed their responses and were given 3 min for recall.

Results

Correct recalls

Consistent with prior research there was a robust internal context change effect in that participants who recalled the words in the same internal context as they were learned recalled more words ($M = .34$, $SE = .03$) than participants who recalled in a different internal context ($M = .24$, $SE = .03$), $t(46) = 2.71$, $p < .01$, $\eta^2 = .14$. As shown in Fig. 4a, examining proportion correct as a function of serial position suggested a main effect of serial position, $F(9, 414) = 12.29$, $MSE = .06$, $p < .01$, partial $\eta^2 = .21$, such that there was a strong primacy effect but no recency effect. Note for analyses we combined every four serial positions. Similar to Experiment 1 the groups demonstrated equivalent serial position effects and there was no context group by serial position interaction, $F(9, 414) = .95$, $MSE = .06$, $p > .48$, partial $\eta^2 = .02$. Again, although there was a main effect of context, the two groups demonstrated largely similar serial position functions.

Similar to Experiment 1 we next examined how participants initiated recall and transitioned to new items during the recall period by examining PFR and lag-CRP functions. Shown in Fig. 4b are the PFR functions for both conditions. As can be seen there was a main effect of serial position, $F(9, 414) = 8.33$, $MSE = .08$, $p < .01$, partial $\eta^2 = .15$, suggesting that participants primarily initiated recall with primacy items. Furthermore, as shown in Fig. 4b, the PFR functions did not differ for context change and context same conditions, $F(9, 414) = .39$, $MSE = .08$, $p > .94$, partial $\eta^2 = .01$, suggesting that participants in both conditions initiated recall in a similar fashion. Shown in Fig. 4c are the lag-CRP functions for forward and backward transitions. Consistent with Experiment 1 and previous research (Howard & Kahana, 1999; Kahana, 1996) the majority of transitions were of a lag of 1 and in the forward direction. Specifically, transitions associated with a short lag were more likely than transitions associated with a long lag, $F(4, 184) = 24.43$, $MSE = .004$, $p < .01$, partial $\eta^2 = .35$, and the lag effect was stronger in the forward than backward direction, $F(4, 184) = 3.41$, $MSE = .005$, $p < .01$, partial $\eta^2 = .08$. Thus, participants tended to begin recall with the first word presented in a list and then tended to recall items in the forward direction leading to large primacy and virtually no recency effects. As shown in Fig. 4c, there were no differences between the context conditions in terms of the lag-CRP functions. Specifically, neither the lag by condition interaction nor the direction by lag by condition interaction reached conventional levels of significance, all p's > .56. Thus, despite the fact that participants in the context same condition recalled more items than participants in the context change condition, these groups were strikingly similar in terms of how they initiated recall and transitioned between correct recalls.

Error responses

Error responses were examined next. Similar to Experiment 1errors were classified as intrusions or repetitions. Overall the two conditions made similar numbers of errors. Specifically, although participants in the context change condition had numerically more intrusions ($M = 2.33$, $SE = .46$) than participants in the context same condition ($M = 1.46$, $SE = .39$) this difference did not reach conventional levels of significance, $t(46) = 1.45$, $p > .15$, $\eta^2 = .04$. Additionally, there was no difference in repetitions between the context same ($M = .54$, $SE = .16$) and the context change conditions ($M = .29$, $SE = .09$), $t(46) = 1.35$, $p > .18$, $\eta^2 = .04$. Thus, the changes in context did not influence the number or type of errors generated.

Latency measures

We also examined latency and IRT information to better understand the dynamics of delayed free recall. Shown in Fig. 5 are the cumulative recall curves for the context same and context change conditions. Consistent with previous research (Rohrer & Wixted, 1994; Wixted & Rohrer, 1994), the cumulative recall curve was well described by a cumulative exponential. As shown in Fig. 5 the symbols represent the data and the lines represent the best fitting cumulative exponential. The resulting parameter estimates suggest that the context same and context change conditions differed in N, but did not seem to differ in λ. Specifically, for the context same condition the parameter estimates were $N = 16.13$ and $\lambda = .10$. For the context change condition the parameter estimates were $N = 12.28$ and $\lambda = .11$. Additionally, the fits were acceptable with the functions each accounting for 99% of the variance and Kolmogorov–Smirnov tests were non-significant ($p > .72$). Thus, the cumulative recall functions suggest that the two conditions differ in the total number of items recalled but the rate of approach to asymptotic levels did not differ.

Next, we examined recall latency in more detail by examining the time to the first response, average recall latency, and average IRTs. As shown in Table 2, and similar to Experiment 1, there were no differences between the two context conditions in terms of the recall latency measures. Specifically, in both conditions the first item was emitted around 5–6 s after the onset of the recall signal and there was no difference between the conditions, $t(46) = 1.06$, $p > .30$, $\eta^2 = .02$. Overall, average recall latency was approximately 37 s and this did not differ significantly between conditions, $t(46) = .14$, $p > .89$, $\eta^2 = .00$. Finally, overall IRTs were approximately 7–8.5 s and this also did not differ significantly as a function of condition, $t(46) = 1.44$, $p > .15$, $\eta^2 = .04$. As with Experiment 1, these results suggest that although there were differences in the total number of items recalled in the two conditions, there were no differences in the dynamics of recall as measured by various recall latency measures. Thus, the time to output responses was the same in the two conditions.

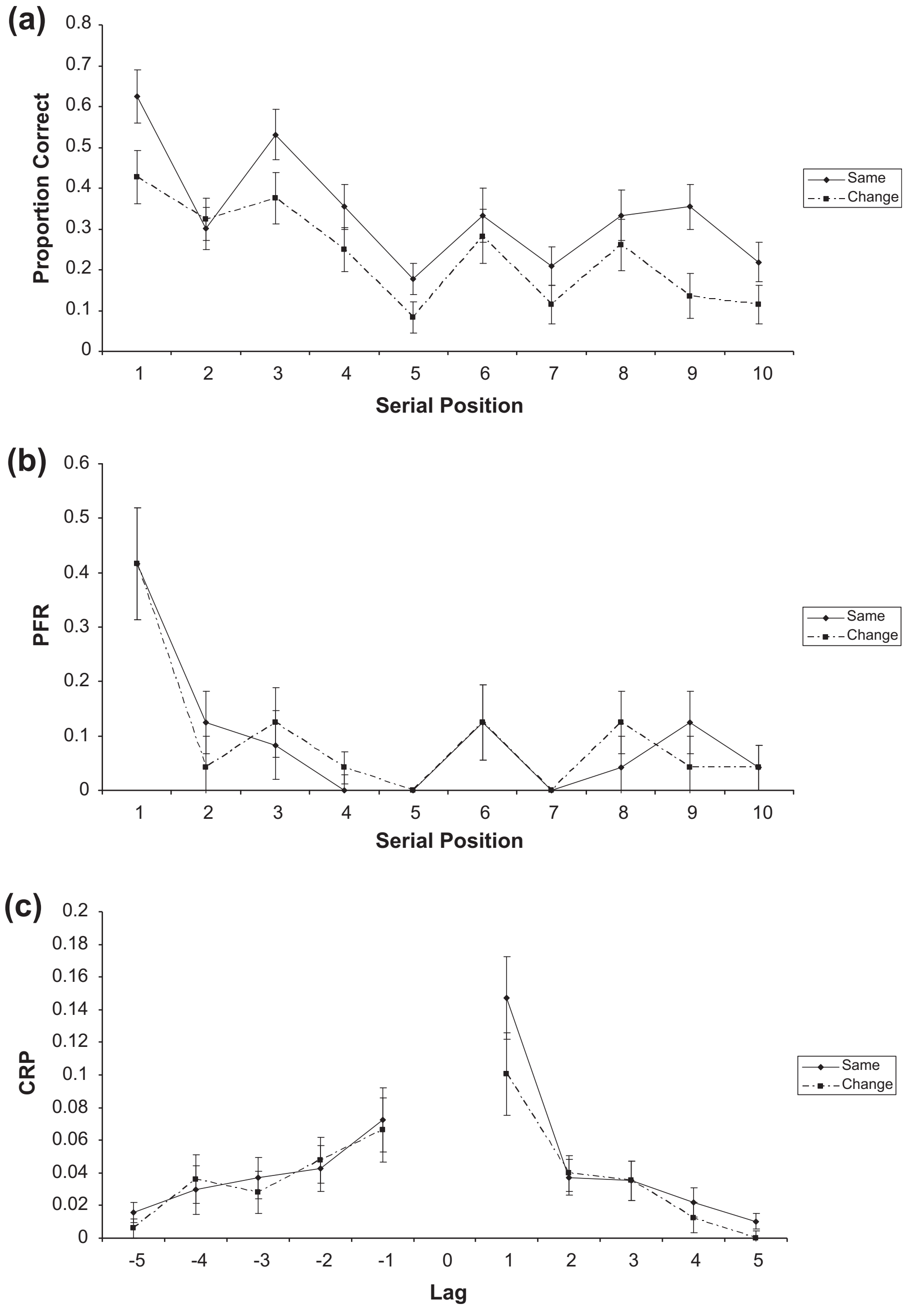

Fig. 4. (a) Proportion of correct recall as a function of serial position and context condition. (b) Probability of first recall (PFR) as a function of serial position and context condition. (c) Conditional response probability functions for forward and backward transitions per list as a function of lag and context condition. Error bars represent one standard error of the mean.

Discussion

The results from Experiment 2 were remarkably similar to those from Experiment 1. Specifically, a putative change in internal context (e.g., Delaney et al., 2010; Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002) resulted in fewer items being recalled than when context remained the same. Despite differences in overall levels of recall the change in context did not change the serial position function, the PFR function, or the lag-CRP function suggesting that participants regardless of condition largely recalled primacy items, recalled primacy items first, and transitioned in a forward manner to temporally close items. Similar to Experiment 1, the conditions did not differ on any of the recall latency measures. Thus, the way in which participants recalled items and the rate at which they recalled

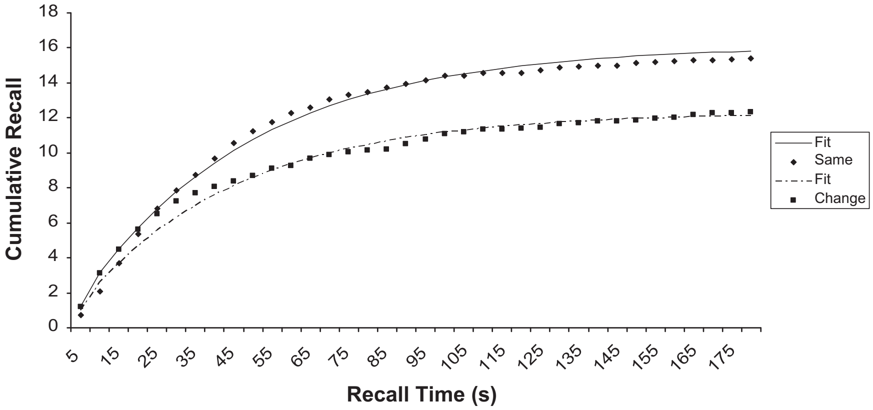

Fig. 5. Cumulative recall curves as a function of recall time and context condition. Symbols represent the observed data and the solid line represents the best fitting exponential.

Table 2
Recall latency measures (in s) as a function of context condition in Experiment 2.

Condition	Measure		
	Time-to-first	Recall latency	IRT
Same	6.07 (.66)	37.30 (3.04)	7.05 (.63)
Change	5.04 (.72)	37.93 (3.54)	8.50 (.79)

Note: IRT = inter-response time. Standard errors are shown in parentheses.

items was the same despite the change in context. Like Experiment 1, these results are most consistent with the notion that changes in context affect the recoverability of items, but do not affect the size of the search set or affect the way in which participants output items. Given these overall similarities between the experiments, these results further suggest that external and internal changes in context largely affect the same cognitive mechanisms leading to similar patterns of results.

Cross-experimental analyses

Given the similarity between Experiments 1 and 2, we further examined the data via cross-experimental analyses. This was done in order to better examine IRTs across the experiments and to better examine how the data matches the predicted cumulative recall curves presented previously. In terms of IRTs we wanted to examine a prediction of the basic search model presented previously. Specifically, when recall latency is constant across conditions, but more items are recalled in one condition than another, it follows that the condition with the fewest items recalled should have a longer mean IRT than the condition where more items are recalled. For example, according to the random search model if two conditions have mean latencies of 35 s and in one condition 12 items are recalled whereas in another 15 items are recalled, then the mean IRT for the condition with fewer recalls should be longer than the mean IRT associated with more recalls (i.e., 9.61 s vs. 8.13 s; see Eq. (4) of Rohrer, 1996).[2] Thus, in the current experiments this would mean that because fewer items were recalled in the context change condition than the context same condition and even though recall latency was the same, mean IRTs should have been longer in the context change condition than in the context same condition. Looking at Tables 1 and 2 suggests that in both cases there was trend in the expected direction, but it was not significant. To better examine this, and to demonstrate that the timing measures were sensitive to the manipulations, we combined the data for Experiments 1 and 2 and reanalyzed mean IRT differences between the conditions. Consistent with the predictions of the random search model, mean IRT in the context change condition was longer ($M = 9.17$, $SE = .54$) than mean IRT in the context same condition ($M = 7.63$, $SE = .49$), $t(84) = 2.12$, $p < .05$, $\eta^2 = .05$. Thus, although there were no differences between the conditions in overall recall latency, there were differences in mean IRT as predicted by the random search model suggesting that the timing measures in the current study were, in fact, sensitive to the manipulations.[3]

Next we wanted to examine the extent to which the data (collapsed across experiments) matched the predicted cumulative recall curves presented in Fig. 1. Recall that we generated predictions for a context same condition along with predictions for three different context change possibilities that suggested that context either changes the recoverability of items (labeled Nonrec in the figure), increases the number of irrelevant items in the search set (labeled Large in the figure), or decreased the number of targets in the search set (labeled Small in the figure).

[2] We thank John Wixted for pointing this out to us and for suggesting these analyses.

[3] Not only does the random search model predict that mean IRT should differ between the conditions, but it also predicts that the last IRT for both conditions should be equal and the last IRT should equal overall mean recall latency (see Eq. (4) of Rohrer, 1996). This was indeed the case in the current study. The last IRT did not differ between the two context conditions (M same = 31.96 s, $SE = 3.34$ vs. M change = 33.96 s, $SE = 3.46$), $t(84) = .41$, $p > .68$. Furthermore, the last IRT did not differ significantly from overall mean recall latency, $t(85) = 1.60$, $p > .11$.

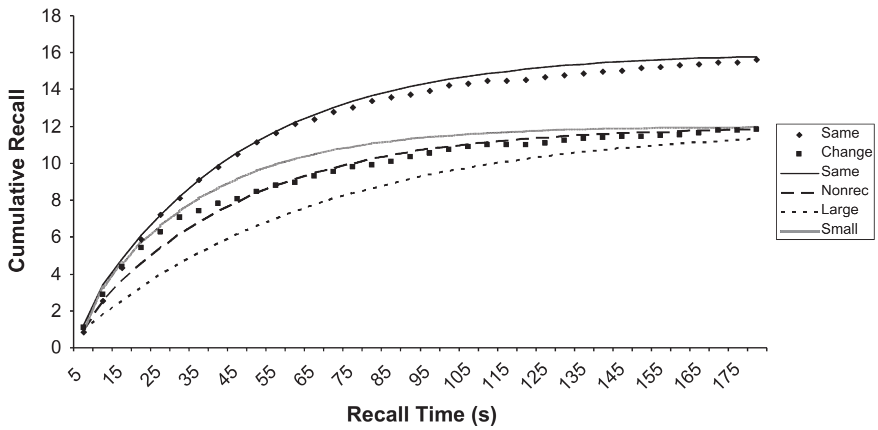

Fig. 6. Comparison of predicted cumulative recall functions with actual cumulative recall functions (average of Experiments 1 and 2). Symbols reflect actual data and the lines reflect the four different predicted cumulative recall functions. See text for details.

Shown in Fig. 6 are the four predicted curves from Fig. 1 along with the combined data from Experiments 1 and 2 for each context condition. The symbols reflect the data and the lines are the four predicted curves. Note, we did not specifically fit the predictions to the data because we were interested in the qualitative pattern of results rather than actual fits to the data. As can been seen in the figure, the predictions from the Nonrec prediction are in line with the actual data, whereas the other predictions (Large and Small) clearly do not match the data in terms of the rate of approach to asymptotic levels. Specifically, the Large prediction condition rises much too slowly and only at the end of the recall period does it start to merge with the data. Conversely, the Small prediction condition rises far too quickly and only merges with the data towards the end of the recall period. These results are consistent with the notion that changes in context in the current study resulted in changes in the recoverability of the items, but did not change the size of the search set.

General discussion

In two experiments changes in context (both external and internal context changes) were examined in free recall. Across both experiments it was found that more items were recalled when encoding and retrieval contexts matched than when they mismatched, thus replicating prior work. Examination of output patterns of recall suggested that in both the context same and context change conditions, participants primarily recalled primacy items, started recall with primacy items, and recalled predominantly in a forward manner resulting in similar serial position functions, probability of first recall functions, and lag-CRP functions. Furthermore, across both experiments, examination of rate of recall suggested that in both conditions participants started recall at the same time, and had similar overall recall latencies. Thus, differences emerged in the proportion of items recalled, but not in how those items were recalled or in overall recall latency. This pattern of results was the same regardless of whether context was changed externally or internally, suggesting that external and internal changes in context affect the same cognitive mechanisms (at least within the current study).

As noted previously, this pattern of results is most consistent with the notion that changes in context influence the recoverability of items. That is, changes in context lead to reductions in associative strengths of items because there are fewer overlapping contextual features between encoded features and features present at test resulting in lowered recovery probabilities (Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980). Similar to changes in presentation duration this means that recovery probabilities should be reduced, but sampling probabilities remain unchanged (Gillund & Shiffrin, 1984). The net effect of this is that fewer items should be recalled, but there should be no differences between the conditions in terms of the rate of recall, and average recall latency. This is exactly the pattern of results that was found, suggesting that in the current study changes in context primarily resulted in changes in the recovery of items.

Other potential mechanisms of context change are possible, but these other mechanism should have resulted in differences in the recall latency variables. For instance, it is also possible that changes in context can also lead to reductions in recall via the introduction of irrelevant information into the search set (Gillund & Shiffrin, 1984; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980; Spear & Riccio, 1994). The inclusion of irrelevant information into the search set would reduce overall sampling probabilities and reduce the likelihood of sampling target items. Much like increases in list-length and increases in proactive interference, this would lead to a lower proportion of items being recalled and overall longer recall latencies because more items are included in the search set (Rohrer & Wixted, 1994; Unsworth, 2007; Wixted & Rohrer, 1993). The fact that there were no differences in recall latency suggest that the size of the search sets for the two conditions was equivalent, thus arguing against this possibility. Likewise the fact that there were no differences in recall latency also argues against the possibility that when context changes fewer target items are actually included in the search set compared to when context is not changed (e.g., Smith, 1994). This would lead to a reduction in recall because not all of the target items are included in the search set and thus, are not even sampled. If this were the case, then one would expect recall latency to have been shorter in the context change condition compared to the context same condition. Again, however this was not the case. Finally, another possibility explored in the current study (and one not mutually exclusive with the other possibilities) is that perhaps changes in context result in changes in how participants use prior recalled items as cues in the next retrieval attempt. That is, perhaps changes in context reduce contextual bindings between items (or changes in the overall retrieval plan) such that once an item is recalled it cannot be used effectively as a cue for the next item. If this were the case, one would expect the context change and context same conditions to have different lag-CRP functions such that in the context change condition the functions should be relatively flat or substantially reduced. However, in both experiments the lag-CRP functions for the two conditions were nearly identical. Thus, it would seem that external and internal changes in context lead to reductions in levels of recall because there are fewer overlapping contextual features between encoded features and features present at test resulting in lowered recovery probabilities.

At first glance, the results of the current studies might seem unsurprising given that many would assume that changes in context would result in fewer overlapping features and lowered recovery probabilities and that the other potential mechanisms we explored were simply straw men. However, there is reason to think that putative changes in context in other paradigms arise from mechanisms other than (or in addition to) changes in recovery. For instance, take the context change account of directed forgetting (Sahakyan & Kelley, 2002) discussed previously. Recall that in list-method directed forgetting studies, participants are presented with a list and in the forget condition participants are told to forget that list and remember the second list. In the remember condition, participants are told to remember both lists. According to the context change account, participants recall fewer items in the forget condition because they intentionally changed their internal context. Like the current study this could be due to in changes in recovery probabilities or changes in sampling probabilities due to the introduction of irrelevant items (i.e., the second list). To test this we (Spillers & Unsworth, 2011) recently had participants perform a standard list-method directed forgetting task and we examined recall latency as well as proportion correct. We found standard costs of directed forgetting such that participants recalled fewer items in the forget condition than the remember condition. In terms of recall latency we found that overall recall latency was longer in the forget condition than in the remember condition. We also found that participants in the forget condition recalled more List 2 items during List 1 recall (i.e., intrusions) than participants in the remember condition. Both of these results are consistent with the notion that recall levels were reduced in the forget condition because List 2 items were included in the search set leading to reductions in sampling probabilities. Thus, if it is assumed that directed forgetting costs are the result of changes in internal context, we must assume that this change in context is somewhat different from changes in context in the present experiments. In one case, recall latency is not changed suggesting differences in recovery, while in the other case recall latency is increased suggesting the inclusion of irrelevant information (junk memoires) in the search set.

The key difference between the current experiments and directed forgetting studies is the fact that in directed forgetting a second list is always presented after the target list (Pastötter & Bäuml, 2007). The presentation of the second list results in the inclusion of irrelevant information into the search set leading to reductions in levels of recall and increases in recall latency. Indeed, Pastötter and Bäuml (2007) recently suggested that a second list was necessary to find directed forgetting costs and context-dependent forgetting. Clearly, the current results suggest that a second list is not necessary in order to find context-dependent forgetting, but it should certainly increase the likelihood of finding context-dependent forgetting because not only would recovery probabilities be reduced, but sampling probabilities would also be reduced. Thus, putative changes in context can arise from differences in recovery (due to fewer overlapping contextual features) as well as differences in sampling probabilities (due to the inclusion of irrelevant items into the search set). When a single list is presented and context is changed (either internally or externally) reductions in recall are likely due to changes in recovery. When two (or more) lists are presented and context is changed reductions in recall are likely due to the inclusion of irrelevant information in the search set (as well as possibly changes in recovery). Examinations of recall latency in addition to traditional measures of proportion recalled can be useful in elucidating which mechanisms are the primary reasons for forgetting.

The current experiments demonstrated robust context-dependent memory effects in free recall. The examination of multiple dependent measures in terms of search models of free recall suggest that the locus of this effect resides in differences in the recoverability of items. However, other potential mechanisms of context-dependent recall are possible, especially when more than one list is presented. Furthermore, it is possible that more elaborate search models that allow for variable items strengths, inter-item associations, and strategic search processes like search termination rules could provide a different interpretation of the current results. Future work is needed to better explore how context change manipulations influence various measures of performance and what potential mechanisms are influenced by various manipulations of context change.

As the opening quotes suggest, context plays an especially prominent role in theories of episodic memory and despite much theorizing, the ways in which context influences memory are still not fully understood. Recent advances in understanding recall dynamics and output dynamics in free recall can be used to shed light on these reliable and important effects which have been present since the early days of experimental psychology.

Acknowledgment

Thanks to Kenneth Malmberg, Jeroen Raaijmakers, and John Wixted for comments of on a previous version of the article.

References

Abernethy, E. M. (1940). The effects of changed environmental conditions upon the results of college examinations. *Journal of Psychology, 10*, 293–301.
Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.
Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of restricted associative responses. *Journal of General Psychology, 30*, 149–165.
Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129–148.
Bower, G. H., Monteiro, K. P., & Gilligan, S. G. (1978). Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior, 17*, 573–585.
Carr, H. A. (1925). *Psychology: A study of mental activity*. New York: Longmans Green.
Delaney, P. R., Sahakyan, L., Kelley, C. M., & Zimmerman, C. A. (2010). Remembering to forget: The amnesic effect of daydreaming. *Psychological Science, 21*, 1036–1042.
Eich, J. E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition, 8*, 157–173.
Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62*, 145–154.
Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1080 words. *Behavior Research Methods and Instruments, 14*, 375–399.
Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.
Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351–360.
Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: Land and underwater. *British Journal of Psychology, 66*, 325–331.
Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple trace memory model. *Psychological Review, 95*, 528–551.
Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 923–941.
Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*, 269–299.
Indow, T., & Togano, K. (1970). On retrieving sequences from long-term memory. *Psychological Review, 77*, 317–331.
Isarida, T., & Isarida, T. K. (2006). Influences of environmental context on the recency effect in free recall. *Memory & Cognition, 34*, 787–794.
Isarida, T., & Isarida, T. K. (2007). Environmental context effects of background color in free recall. *Memory & Cognition, 35*, 1620–1629.
Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*, 103–109.
Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 530–540.
Klein, K., Shiffrin, R. M., & Criss, A. H. (2007). Putting context into context. In J. S. Nairne (Ed.). *The foundations of remembering: Essays in honor of Henry L. Roediger* (vol. 3, pp. 171–190). Psychology Press.
McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352–370.
McGill, W. J. (1963). Stochastic latency mechanism. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.). *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York: Wiley.
Mensink, G., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review, 95*, 434–455.
Nixon, S. J., & Kanak, N. J. (1981). The interactive effects of instructional set and environmental context changes on the serial position effect. *Bulletin of the Psychonomic Society, 18*, 237–240.
Pastötter, B., & Bäuml, K. H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 977–982.
Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. Bower (Ed.). *The psychology of learning and motivation* (Vol. 14). New York: Academic Press.
Rand, G., & Wapner, S. (1967). Postural status as a factor in memory. *Journal of Verbal Learning and Verbal Behavior, 6*, 268–271.
Roediger, H. L., III, Stellon, C. C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning & Memory, 3*, 174–188.
Rohrer, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition, 24*, 188–201.
Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition, 22*, 511–524.
Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 1064–1072.
Saufley, W. H., Otaka, S. R., & Bavaresco, J. L. (1985). Context effects: Classroom tests and context independence. *Memory & Cognition, 13*, 522–528.
Schab, F. R. (1990). Odors and remembrance of things past. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 648–655.
Shiffrin, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 375–447). New York: Academic Press.
Smith, S. M. (1985). Background music and context-dependent memory. *American Journal of Psychology, 98*, 591–603.
Smith, S. M. (1994). Theoretical principles of context-dependent memory. In P. Morris & M. Gruneberg (Eds.), *Aspects of memory* (2nd ed., pp. 168–195). Routledge Press.
Smith, S. M., Glenberg, A. M., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*, 342–353.
Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*, 203–220.
Spear, N. E., & Riccio, D. C. (1994). *Memory: Phenomena and principles*. Boston: Allyn & Bacon.
Spillers, G. J., & Unsworth, N. (2011). Are the costs of directed forgetting due to failures of sampling or recovery? Exploring the dynamics of recall in list-method directed forgetting. *Memory & Cognition, 39*, 403–411.
Strand, B. Z. (1970). Change of context and retroactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 9*, 202–206.
Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1020–1034.
Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language, 59*, 223–236.
Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition, 37*, 837–849.
Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology, 31*, 1–23.
Weir, W., & May, R. B. (1988). Environmental context and student performance. *Canadian Journal of Education, 13*, 505–510.
Wixted, J. T., Ghadisha, H., & Vera, R. (1997). Recall latency following pure-and mixed-strength lists: A direct test of the relative strength model of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 523–538.
Wixted, J. T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1024–1039.
Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An intergrative review of the empirical literature. *Psychonomic Bulletin & Review, 1*, 89–106.