A Novelty-Induced Change in Episodic (NICE) Context Account of Primacy Effects in Free Recall

Eddy J. Davelaar

Department of Psychological Sciences, Birkbeck, University of London, London, UK
Email: e.davelaar@bbk.ac.uk

Copyright © 2013 Eddy J. Davelaar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Formal cognitive models of episodic memory assume that during encoding list items become associated with a changing context representation. However, this representation is recency-biased and thus can not account for primacy effects under conditions that prevent rehearsal. In this paper, it is hypothesized that one source underlying primacy effects is the detection of novelty. In three experiments, it is shown how novelty at the perceptual and semantic level can explain the full serial position function of first recall probabilities, including primacy effects. It is proposed that an item becomes distinctive due to increase in the change within a distributed episodic context representation, induced by novelty detection. The theory makes three assumptions. First, items become associated with a distributed context representation. Second, the context representation changes with item presentation. Third, the rate of contextual change is related to the perceptual and conceptual difference computed between the presented item and the previous item (or items in the buffer). This theory captures primacy effects in first recall probabilities without recourse to a rehearsal process and provides a mechanistic account of distinctiveness.

Keywords: Distinctiveness; Novelty-Detection; Episodic Context; Distributed Context; Contextual Change

Introduction

One of the most robust results in cognitive psychology is the U-shaped serial position curve obtained in the immediate free recall task (Murdock, 1962). In the free recall task, participants are presented with a sequence of words and are instructed to recall as many words as possible in any order. Words presented at the beginning or end of the sequence are remembered better than words presented in the middle of the sequence. These phenomena are referred to as primacy and recency effects, respectively. There exists a number of theories that can account for these results and have advanced our understanding of the mechanisms of human memory.

Primacy effects in list memory have been explained in a variety of ways. Many dual-store and dual-trace models assume that the first item enters an empty short-term memory (STM) buffer from which it is displaced after the buffer is filled to capacity, leading to that item residing in STM longer than subsequent items. As episodic trace strength is a function of the duration that items reside in the buffer, the episodic trace for the first item will be stronger than subsequent items (Atkinson & Shiffrin, 1968; Davelaar et al., 2005; Raaijmakers & Shiffrin, 1980). Other encoding models explain primacy effects in terms of more opportunity for rehearsing the first items (Murdock & Metcalfe, 1984; Tan & Ward, 2001; Ward, 2002). There are two types of rehearsal explanations. In one, increased rehearsal opportunities for the first few items lead to increased opportunities for encoding and hence to a stronger trace. This rehearsal-enhanced encoding is assumed in early work on re-

hearsal (Atkinson & Shiffrin, 1968; Rundus, 1971; Rundus & Atkinson, 1971; Brodie & Murdock, 1977). In the second type of rehearsal explanation, increased rehearsal opportunities for the first few items lead to increased probabilities that the items are still in the rehearsal cue at the end of the list presentation. This rehearsal-enhanced *accessibility* of the items is assumed in the more recent work on rehearsal (Tan & Ward, 2001) in which no short-term buffer is presumed. This denial of a short-term buffer requires that such models liken the rehearsal process that occurs during list presentation to mini-retrievals (Ward, 2002; see also Laming, 2006). Retrieval-based explanations of primacy effects appeal to the notion of the first items being more distinctive (Murdock, 1960; Neath, 1993). This notion has proven useful to account for the finding that after retrieval of end-of-list items, participants continue retrieving begin-of-list items, as if *the first item contains a tag that is accessible at retrieval* (Davelaar et al., 2005; Murdock & Metcalfe, 1984). Although the explanations in terms of buffer-enhanced encoding, rehearsal-enhanced encoding, rehearsal-enhanced accessibility, and distinctiveness all capture the basic pattern, they are not mutually exclusive.

When procedures are in place that minimizes the use of rehearsal, primacy is still found. For example, Richardson and Baddeley (1976) found primacy effects in immediate free recall under articulatory suppression. Primacy effects in immediate free recall are also found in an incidental task (Baddeley & Hitch, 1977) and when participants are required to make a semantic judgment for each word (Howard & Kahana, 1999). In a procedure called the continuous-distractor paradigm, each item

is preceded and followed by an interval of distractor activity. Even though this procedure abolishes the opportunity to rehearse, the items during the distractor activity and equates the duration that each item resides in the buffer, primacy effects are still found (e.g., Bjork & Whitten, 1974; Neath, 1993; Thapar & Greene, 1993). These results suggest that not all primacy gradients found in serial position functions are a consequence of a rehearsal mechanism or a buffer. However, this is not to say that *all* primacy gradients are due to distinctiveness-based retrieval mechanisms.

Evidence supporting the view that primacy effects may have an origin during encoding comes from neuroscientific investigations. Electrophysiological studies have revealed that enhanced gamma band activity is correlated with enhanced memory performance for earlier list positions (Sederberg et al., 2006). In the neuroscientific literature these results have been discussed with reference to the computation of novelty (Debener, Herrman, Kranczioch, Gembris, & Engel, 2003): the higher the novelty of an item, the larger the neural response. The neural response for the first item would then be the result of a larger difference between the first item and the fixation stimulus (that is typically presented before the first item) than between the second and the first item. Importantly, it has been shown that increased activation in the same brain areas that compute novelty also predicts recall performance (Kirchoff et al., 2000).

Whereas the neuroscientific evidence points to the medial-temporal cortex as the site for novelty signals that are related to increased memory, neurocomputational models of memory have implicated the medial-temporal cortex as a contextual system (e.g., Howard et al., 2005; Norman & O'Reilly, 2003). This triple conjunction of novelty, contextual change and enhanced memory in the same brain area raises the question whether a novelty signal is directly related to contextual change, which in turn is related to increased memory performance. In what follows next, I will put forward the hypothesis that novelty is one source for change in episodic context and that contextual change leads to enhanced memory for the first item after the change. The hypothesis is critically different from its nearest neighbors. That is, the novelty signal itself does not lead to an enhanced memory trace or causes an increase in dimensional distinctiveness at retrieval. Instead, the contextual change at encoding leads to the first item after a change to be relatively less different from the context during retrieval than expected based on temporal distance alone. As above, the proposed hypothesis is compatible with the other explanations of primacy effects and the question which view is superior is not taken up here.

Together, the behavioral data on primacy effects supports the existence of a mechanism that favors the first item in a list and the neuroscientific data points towards a novelty-signal that is computed on-line. To directly test the idea that novelty contributes to primacy effects, a series of experiments was conducted addressing the following questions. First, can primacy effects be enhanced by making the first item in a list more novel? Second, can novelty during list presentation create a midlist primacy effect? To preview, the data confirms the role of novelty in generating primacy effects. In the next section, a formalization of the novelty-induced primacy effect is presented in what will be called the novelty-induced change in episodic context model or *NICE* context model for short. This is then followed by three experiments and a general discussion.

Distinctiveness through Novelty-Driven Change in Episodic Context

Within the memory literature, the term *distinctiveness* has been used as an explanation for such well-known empirical findings as the primacy effect in serial position functions (Murdock, 1960; Neath, 1993) and the isolation effect (Brown, Neath, & Chater, 2007). However, some researchers have argued that "Distinctiveness itself cannot be used as an explanation for effects of distinctiveness" (Hunt & Lamb, 2001: p. 1359). Instead *perceived novelty* leads items to become encoded in episodic contexts that have just undergone a relatively large change and are therefore seen at retrieval as being more distinctive. In other words, *distinctiveness is not a process, but an end-product*. The model presented here adds a further nuance in that local distinctiveness and encoding-retrieval match are tightly connected.

In this paper, a novelty-based mechanism is proposed that produces primacy effects. The model is a changing-context model in which contextual change is driven by novelty. The model is named the Novelty-Induced Change in Episodic (NICE) context model and has three critical assumptions, two of which were employed successfully in previous changing-context models (i.e., Glenberg et al., 1983; Howard & Kahana, 1999; Mensink & Raaijmakers, 1988). First, context is a distributed representation of active and inactive elements. Second, the context representation changes slowly over time by making active elements inactive (with probability β) and inactive elements active (with probability α). Third, the probability of activation is a positive function of the novelty of a stimulus. In short, the model is a Markov system with novelty-dependent transition probabilities. Each of these assumptions will be addressed in turn.

Distributed Context

In following several investigators (e.g., Burgess & Hitch, 1999; Dennis & Humphreys, 2001; Glenberg et al., 1983; Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Norman & O'Reilly, 2003) context is a distributed representation of active and inactive elements. The anatomical location of this context representation may be debated, but in some computational models the entorhinal cortex in the medial-temporal lobe is assumed to be the location (Howard et al., 2005; Norman & O'Reilly, 2003). It is known that patients with damage to the medial-temporal lobe suffer from long-term memory loss (Scoville & Milner, 1957; Baddeley & Warrington, 1970), being unable to recall items that were presented early in the list. Assuming that the elements correspond to (groups of) neurons in the medial-temporal lobe allows an investigation whether the model is able to address the finding that activation in the medial-temporal lobe reflects not only the amount of novelty detected, but also is predictive of subsequent memory.

Contextual Change

During presentation of a stimulus, active contextual elements may be de-activated, while currently inactive elements may be activated. This process is probabilistic with β being the probability that an active element becomes inactive and with α being the probability that an inactive element becomes active. In following Mensink and Raaijmakers (1988; see also Estes, 1955; Murdock, 1972), the total number of active elements $A(t)$

at time t is given by:

$$A(t) = A(0)e^{-(\alpha+\beta)t} + N\frac{\beta}{\alpha+\beta}\left(1 - e^{-(\alpha+\beta)t}\right)$$

which asymptotes at $N\beta/(\alpha + \beta)$ (the maximum number of active elements). $A(0)$ is the number of active elements at time = 0.

Novelty-Dependency

In the models of Mensink and Raaijmakers (1988) and Howard and Kahana (2002), the transition probabilities were assumed to remain constant. However, this assumption makes the system's dynamics independent of the nature of the items. This counters the evidence reviewed above showing systematic variations in brain activation based on the type of material and more specifically the novelty of an item. Here, it is assumed that β and α are variable (see also Murdock, 1972) and are positively correlated with the novelty of an item. Novelty, γ, is measured as the total number of attended features in a (distributed) stimulus representation that become activated. Stimulus representations contain semantic features of the to-be-remembered word and the features of the background picture on which the word is presented, but only those that are attended to by the participant. The probabilities α and β for a given γ could then be formulated as $\alpha_\gamma = F(\alpha_0, \gamma)$ and $\beta_\gamma = G(\beta_0, \gamma)$. Here, α_0 and β_0 are the minimal probabilities that are present when a stimulus is immediately repeated. The functional form of equations $F(\cdot)$ and $G(\cdot)$ are not known, but suffice is to say that $\beta_\gamma > \beta_0$ for several of the data sets looked at. Future work could address the precise equations in detail. These equations show that the distance between the values of β and α becomes smaller with increase in novelty and implies that the asymptotic level of the total number of active elements, $N\beta/(\alpha + \beta)$, would then become larger (with $N/2$ as its maximum). This is important, as it suggests that with these three assumptions in place, the number of active elements that will be encoded with the stimulus is larger for a novel than for a similar stimulus. In other words, similar stimuli are encoded in similar contexts. Novel stimuli are encoded in dissimilar contexts (increased contextual change) with more contextual elements.

Memory Encoding and Retrieval

During encoding, active stimulus features and active contextual elements become associated. This matrix of connections represents the episodic memory of the events and is used in the retrieval of those events. Memory performance is assumed to be dependent on the similarity between the context at encoding and the context at test (Tulving, 1983; see Nairne, 2002 for a challenge). This is equivalent to the number of elements that are active at the time of test and were associated with a particular stimulus. The details of a retrieval process are less important than the actual consequence that novelty-driven contextual change has on the retrieval probability (for detailed models see Howard & Kahana, 2002; Mensink & Raaijmakers, 1988). However, several general aspects of this contextual retrieval are relevant. First, the similarity and therefore the probability of retrieval decreases with increasing distance between the encoded context and the test context, leading to a recency effect (see Howard & Kahana, 2002). Second, due to the transition probabilities being novelty-dependent, a novel stimulus will be associated with more contextual elements and the context representation will differ more strongly from the context representation associated with its predecessor. This would lead to a higher likelihood of retrieving novel stimuli.

Model Simulation

Figure 1 shows a demonstration of the NICE context model. Vectors of 100,000 binary elements were used in the simulation with values $\alpha = 0.8$ and $\beta = 0.01$. A list of 10 items is simulated in which no novelty signal is provided at the beginning of the list, novelty is provided at the beginning, and where a novelty signal is also provided during presentation of the item in position 6. Novelty is implemented as a momentary increase in $\beta = 0.08$. The striking observation is that the NICE context model is able to produce a primacy effect due to change in episodic context as well as showing an enhancement in memory performance for the novel item in position 6. In other words, the model gave rise to primacy via a process of novelty-induced contextual change.

To understand the reason why the model produces this striking pattern, the dissimilarity among the context vectors that were encoded with the ten items and the context vector prior to the first item (labeled "B") and after the last item (labeled "R") are plotted in a multidimensional scaling (MDS) solution in **Figure 2**. The model was rerun 1,000 times for each of the three lists and subjected to MDS analyses, forcing the retrieval vector, R, in the upper right quadrant.

In the simulation without novelty for the first item, the MDS solution forms a horseshoe in two dimensions with B and R at either side. The MDS solution using three dimensions were not significantly worse, but revealed that along the third dimension the B and R are far removed. When the presentation of the first item coincides with increased novelty, the horseshoe pattern changes such that the vector for the first item moves away from the B vector and gets closer to the R vector. In other words, the first item becomes locally more distinctive from the start of the list and more similar to the retrieval context. When the item in the sixth position coincides with increased novelty, the same happens in that the sixth vector gets closer to the R vector. Moreover, two solutions exist with different orientations for the sequence involving the first five items, which implies that the
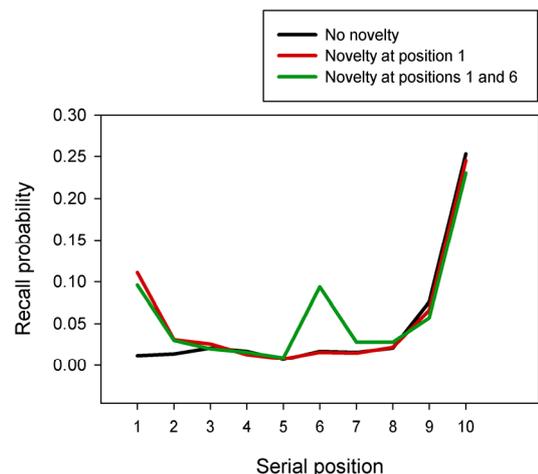

Figure 1.
Simulated serial position functions showing lack of primacy effects without novelty and primacy effects when novelty is present.
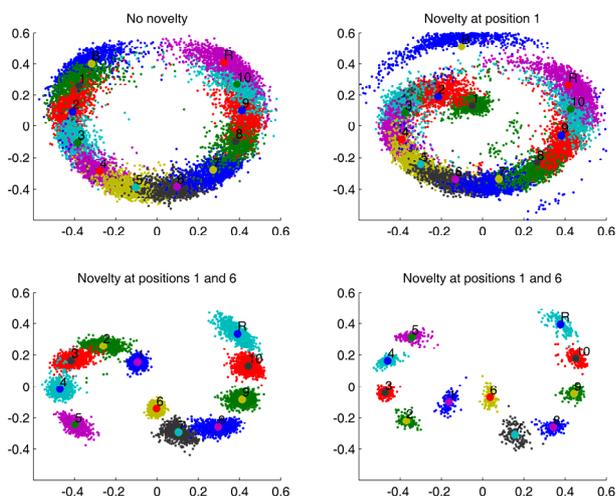
Figure 2.
Multidimensional scaling solutions for the three simulations. The solutions are arranged to have the retrieval context vector, R, in the upper-left quadrant. Each panel includes the context vectors for all ten items and the context vector prior to the first item, B. The lower panels show MDS solutions (without B) of the simulation with novelty in positions 1 and 6. The left and right panels present 840 and 160 solutions, respectively.

novelty during presentation of the sixth item created two sublists. This is further strengthened by the fact that the distance between context vectors 5 and 6 is larger in both solutions compared to the unitary lists. This provides the computational argument that the enhanced memory for the first item after a change is contextually similar to a primacy effect.

Experimental Study

In applying the model to data, the question arises what type of data is suitable. The model only addresses contextual change and therefore can only be tested with data on first recall probabilities. To date no published report addressed first recall probabilities in relation to novelty-induced change. Three free recall experiments were conducted to provide such data. In Experiment 1, the relative novelty of the very first item is manipulated to test the assertion that novelty-induced change in episodic context contributes to primacy effects in the absence of rehearsal. In Experiment 2, the background context is manipulated while participants memorized words. In Experiment 3, a critical item is embedded in a list of category exemplars from a different semantic category.

Experiment 1

In Experiment 1, participants were presented with a modified delayed free recall task in order to explore the mechanisms underpinning the primacy effect. The novelty of the first item was manipulated by preceding it with a variable number of nonword items consisting of random consonants. In addition, we looked at whether rehearsal interacts with novelty. If the primacy effect due to novelty interacts with that induced by rehearsal then it is assumed that both are manifestations of a single process. Thus lack of an interaction can be interpreted as a validation of a separate process that contributes to primacy effects.

The rationale for the particular manipulation of novelty is as

follows. The calculation of novelty implies that a current item is different from the items that preceded it. This difference might be at several levels, such as perceptual (different colors) and conceptual (different semantics). As we used unrelated word in the experiment, each word is equally different conceptually and perceptually from its predecessor, except the very first item. This item is preceded by the fixation prompt that tells the participant that the presentation of the list will begin. Thus, the very first item coincides with additional perceptual change. To equate the differences at the perceptual level, letters are used instead of a fixation cross, such as a plus sign. In addition, it is expected that the computation of novelty is larger after a series of similar items. Thus, in the sequence "+ A B", the word "B" is less novel than in the sequence "+ A A A B". As we are interested in the resulting primacy effect that is due to encoding processes and not due to dynamics unfolding during the retrieval phase, we looked at the first recall of a trial. This is the purest measure of primacy effect, uncontaminated with retrieval-related strategies. We expected that there will be more primacy than recency when there are more nonwords preceding the very first item.

Materials and Methods

Participants

Forty-eight participants (22 females) took part in this experiment. Half were allocated to the no-rehearsal condition and the other half to the rehearsal condition.

Design

The experiment conformed to a $5 \times 10 \times 2$ mixed factorial design crossing the between-subjects factor *rehearsal* ("no rehearsal" group, n = 24, "rehearsal" group, n = 24) with two within-factors: *nonwords* (number of nonwords 0 - 4) and *word serial position*.

Materials

A total of 230 common words and 45 nonwords were used as stimuli. The words were selected such that none had any negative connotations (which could lead to enhanced novelty) and all had a word frequency count between 100 - 1000 in a million. Words within each list did not have any semantic or phonological relation. Letter repetition on the same letter-position was avoided. Nonwords consisted of strings of 3 to 5 consonants, matching the range of word lengths. Example nonwords include: kpnsk, mjn, klbn. Each of the 23 lists consisted of ten words and 0 to 4 nonword items, creating five conditions with 4 lists in each condition (and 3 practice lists).

Procedure

Participants were tested individually and received instructions on the computer screen and verbally by the experimenter. All stimuli were presented on the computer screen. Participants in the rehearsal condition received instructions to rehearse the words from the list during list presentation. After instructions about the delayed free recall task, participants practiced two 10-word lists, one without nonwords (nonword - 0 condition) and one with three nonwords (nonword - 3 condition). The sequence of ten words and variable number of nonwords was presented at a rate of one item per second in the middle of the screen, each item masking the previous one. After the words were presented, a prompt appeared indicating the start of the

distractor task. The distractor task consisted of nine mathematical problems, such as "7 − 3 = 3" and participants indicated the accuracy of the equation by pressing the K- or S-key on the keyboard when it was correct or incorrect, respectively. The maximum duration for each mathematical problem was 2.5 s and participants were advised to respond within 2 seconds to each of the problems as accurately as possible.

Following this distractor activity, three question marks appeared, prompting the participant to start recalling as many words from the list as possible in any order. Although only the very first response was of interest, all answers were record by the experimenter to encourage the participant to memorize all the words during list presentation. After the two practice trials, 21 trials were presented of which the first one was a "warming up list" with two nonwords and was thus excluded from the analyses. The test lasted approximately 25 minutes.

Results and Discussion

Figure 3 presents the serial position curves of the first recall probabilities for the no-rehearsal and the rehearsal group. From the figure it is clear that the distractor task abolished the recency effect in the no-rehearsal group. However, a recency effect is present in the rehearsal group, which is due the fact that participants are able to rehearse items during a difficult distractor task if requested to do so. As the first recall probabilities are normalized (the values sum to one), the rehearsal group necessarily have lower levels of primacy.

As our focus is on the primacy effect, we used the PR-index

Figure 3.
Serial position curves for both groups as a function of the number of nonwords seen prior to the very first item.

(Davelaar et al., 2005) to get a relative estimate of the primacy/recency gradient. The PR-index is larger than 0.5 when there is more recency than primacy and smaller than 0.5 when there is more primacy than recency. Therefore, the PR-index summarizes the entire the serial position curve in a single value. This value varies as a function of the number of nonwords presented as shown in **Figure 4**. The data in **Figure 4** were subjected to a 2 (rehearsal group) × 5 (number of nonwords) mixed ANOVA. There was a main effect of number of nonwords [F(4, 184) = 2.573, MSe = .017, $p < .05$], showing more primacy with increasing number of nonwords in both groups. The main effect of group and the interaction between group and number of nonwords were not significant ($p > .12$).

Participants in both groups showed tendency to recall more primacy words from lists preceded by the maximum four nonword items. This finding supports the hypothesis that the primacy effect is enhanced when the very first item coincides with a larger level of novelty. In addition, instead of enhanced primacy effects in the rehearsal group, the group shows recency effects. This finding shows that the rehearsal process retrieves only the more recent items. Typically, the immediate free recall task is used in which the most recent items can be retrieved from short-term memory and thus rehearsal can not contribute to the performance. However, in the delayed free recall task, all items have been displaced from short-term memory. Any rehearsal process that can unfold will not take the very first item, but instead the most recent item, as expected from a recency-based account of retrieval (Tan & Ward, 2000; Ward, 2002). Finally, the lack of an interaction between novelty and rehearsal confirms that the two are operating on different parts of the list.
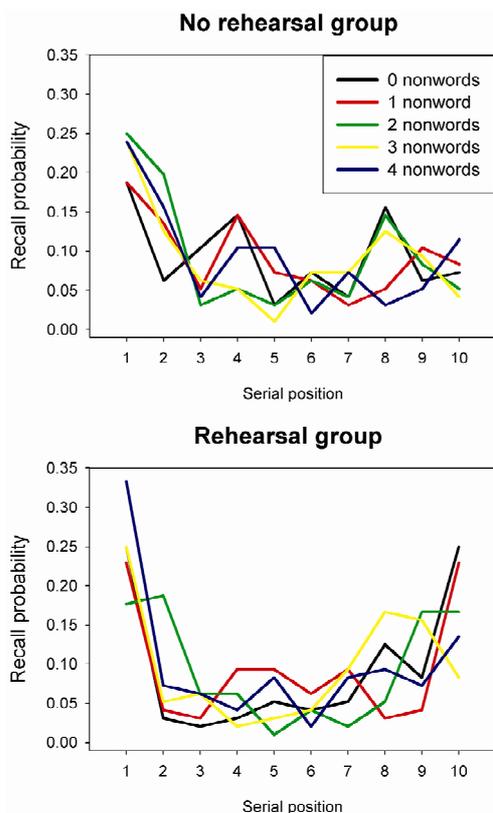
Experiment 2

Experiment 1 confirmed the view that novelty of the very first item contributes to the primacy effect in free recall. This novelty was induced by presenting a number of items that differ from the very first item and were similar to each other. In principle, it should be possible to have a primacy effect present in a single list of items in which the first half of the list differs from the second half of the list. In our conception, the presentation of the very first item of the second half will coincide with a high level of novelty and should thus be better remembered. In Ex-
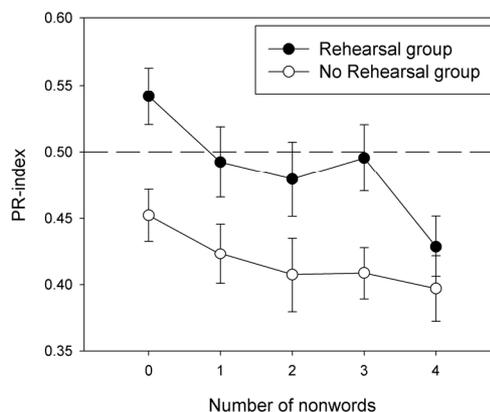

Figure 4.
The primacy/recency index computed over the first recall probabilities. The smaller the index, the more primacy there is in the serial position curve.

periment 2, we varied the background on which words are presented. The rationale for this choice was that the background is unrelated and irrelevant to the task, but forms part of the episodic context in which the word is encoded. When the same context is presented during retrieval, memory for items encoded in that context should be better. This allows investigating whether the match in context at encoding and retrieval interacts with the effect of novelty. If this is the true, novelty and episodic context are related.

Materials and Methods

Participants

Forty-one participants (35 females) took part in this experiment.

Design

The current experiment conformed to a 5 × 10 within-subject design crossing the factors *background* (5 levels) and *word serial position*.

Materials

A total of 150 common words were used as stimuli. The words were selected such that none had any negative connotations and all had a word frequency count between 100 - 1000 in a million. Words within each list did not have any semantic or phonological relation. Letter repetition on the same letter-position was avoided. Each of the 15 lists consisted of ten words. Background pictures were landscapes of forests, beaches, cities, and mountain ranges.

Procedure

Participants were tested individually and received instructions on the computer screen and verbally by the experimenter. Each word was presented for one second on the computer screen on top of a picture that filled the entire screen. Participants were told to ignore the background picture and focus on the words to be remembered. After the final word, participants recalled as many words as possible in any order. There was no time limit imposed for recall.

The background picture varied according to 5 conditions.

Condition 1 (AA/A): the background picture was the same for all 10 words and was also present during the retrieval.

Condition 2 (AA/B): the background picture was the same for all 10 words, but a different picture was presented during retrieval.

Condition 3 (AB/A): words 1-5 were presented on top of a different picture than words 6-10 and during the retrieval the picture that was presented during the first half was re-presented.

Condition 4 (AB/B): same as condition 3, but during retrieval the picture that was presented during the second half was re-presented.

Condition 5 (AB/C): same as condition 3, but during retrieval the picture was a totally different one.

The lists were counterbalanced across participants so that the different lists were presented in each of the different context-conditions. There were 3 lists per condition and no picture was presented in more than one trial.

Results and Discussion

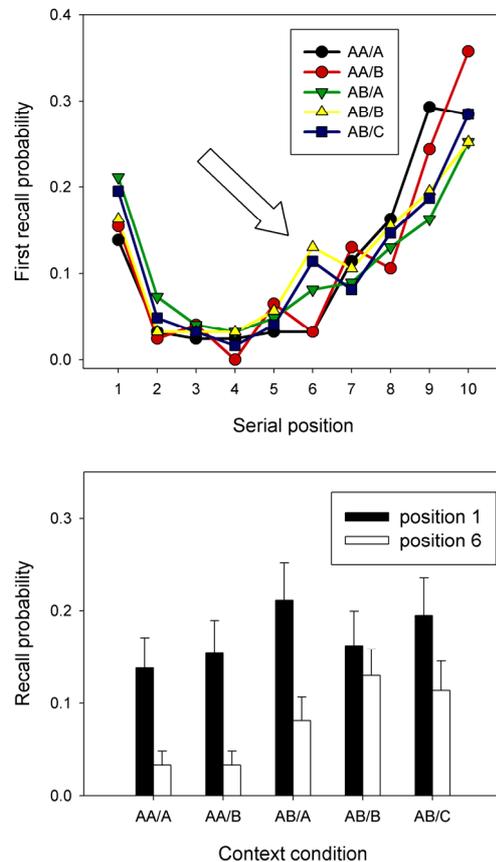**Figure 5** shows the first recall probabilities as a function of

Figure 5.

Top panel: Serial position curves for the conditions in Experiment 2. The first five items were presented with the same background picture, denoted by "A". The second five items were presented with the same background picture, but could be either the same or different than the first list half. During retrieval, a picture stayed on the screen which was either the same or different than the picture presented during encoding. Compare with the simulation results in **Figure 1**. Bottom panel: the raw scores for all experimental conditions.

context condition.

A deviation score was calculated between conditions 2 - 5 and condition 1 (not shown), which serves as a baseline condition, for the words in positions 1 and 6. A 4 (condition) × 2 (list-half) ANOVA conducted on the deviation scores revealed a marginal effect of condition [$F(3, 120) = 2.50$, $MSe = 0.026$, $p = .063$] and a significant contrast effect on the interaction [$F(1, 40) = 6.38$, $MSe = 0.022$, $p < .05$]. This was further unpacked using t-tests. Comparing conditions AA/B and AB/C revealed a significant effect of a contextual change despite the lack of an encoding-retrieval match [$t(40) = 2.23$, $p < .05$]. When the retrieval context matched either list-half, the list position matching the context was enhanced (position 1 in AB/A and position 6 in AB/B), although this did not reach significance. The AB/B [$t(40) = 3.11$, $p < .01$] and AB/C [$t(40) = 2.23$, $p < .05$] conditions showed enhanced recall for position 6. Thus, both change in episodic context during list presentation and contextual encoding-retrieval match contribute to first recall. Finally, comparing the AB/B and AB/C conditions revealed no statistical difference, suggesting that the change-induced en-

hancement and the encoding-retrieval match are not independent.

Experiment 2 reveals that changing an irrelevant background picture is sufficient to produce enhanced recall performance for the first item after the change. In analogy, this constitutes a mini-primacy effect. As shown in the simulations in **Figures 1** and **2**. This is consistent with the view that novelty as induced by change in episodic context contributes to primacy effects. The data in Experiment 2 is too sparse to run correlational analyses relating the size of the primacy effect with the size of the novelty effect. This is taken up in Experiment 3.

Experiment 3

Experiments 1 and 2 provide the first evidence of primacy effects being driven by novelty. However, in Experiment 1, the novelty was induced at the item level, whereas in Experiment 2 this was induced at the level of episodic context. It is therefore possible that the mini-primacy effect was driven by the change in context and not necessarily by item-level novelty. The third and final experiment employs a delayed free recall task of semantically related items in which one item in the middle of the list is semantically unrelated to the rest. The novelty for that item is then based on item-level analysis. This unrelated item, the isolate, is positioned either at the beginning or in the middle of the list. If item-level novelty (and not just context-level novelty) contributes to primacy effects, then normal primacy effects should correlate with the increased memory for the isolate when presented in the middle list position.

Materials and Methods

Participants

Seventeen participants (11 females) took part in this experiment.

Design

The current experiment conformed to a 2 × 12 within-subject design crossing the factors *isolate position* (position 1 or position 6) and *word serial position*.

Materials

A total of 240 common words were used as stimuli. The words were selected such that none had any negative connotations and all had a word frequency count between 100 - 1000 in a million. Words within each list were semantically related by virtue of being exemplars for a single category. Each of the 20 lists consisted of twelve words. A word that was unrelated to the words in the list was presented either on position 1 (b-list) or position 6 (m-list). There were 10 lists per condition and the location of the isolate was counterbalanced across participants.

Procedure

Participants were tested individually and received instructions on the computer screen and verbally by the experimenter. Each word was presented for one second on the computer screen. After the final word, a prompt appeared indicating the start of the distractor task. The distractor task consisted of nine mathematical problems, such as "7 − 3 = 3" and participants indicated the accuracy of the equation by pressing the K- or S-key on the keyboard when it was correct or incorrect, respectively. The maximum duration for each mathematical problem

was 2.5 s and participants were advised to respond within 2 seconds to each of the problems as accurately as possible. Following this distractor activity, three question marks appeared, prompting the participant to start recalling as many words from the list as possible in any order. There was no time limit imposed for recall.

Results and Discussion

Figure 6 presents the first recall serial position curves. As can be seen, the isolate is reported more often that the words in the same position, but related to the other words in the list. This constitutes the basic isolation effect also known as the Von Restorff effect. Our focus is on whether the size of the isolation effect on position 6 (m-list minus b-list) correlates with the normal primacy effect (position 1 in the m-list) and with the additional benefit when the isolate is in position 1 (b-list minus m-list). To provide a baseline, the correlational analyses were repeated for all other serial positions (2 - 5, 7 - 12). The correlation between the recall difference in position 6 correlated significantly with the primacy effect [$r(17) = -.63$, $p < .01$] and with the addition to the primacy effect [$r(17) = .91$, $p < .001$]. The only other position that showed significance was the final list position with the normal primacy effect [$r(17) = -.51$, $p < .05$]. This pattern makes sense given that for the recall, the end-of-list context is used as a memory cue and is directly responsible for the first recalls. The last item is encoded in an episodic context that is similar to the end-of-list context. However, the final list item does not coincide with enhanced novelty and therefore it does not correlate with the novelty-induced enhancement on the primacy effect or even with the novel item itself.

The results support the view that item-level novelty drives better first recall performance of the items and that normal primacy effects is partly driven and is enhanced by item-level novelty. As an aside, the lack of correlations between the primacy effects and item number 7 suggests that the semantic novelty of item 7 is not as great as item number 6. This suggests that novelty is calculated against all items in short-term memory, which is consistent with the results of Experiment 1 where the primacy gradient is gradually enhanced when placing more nonwords in short-term memory. However, Experiment 2 clearly showed a close connection between novelty-induced memory enhancement and change in episodic context. The overall pattern of results can be accommodated in the NICE
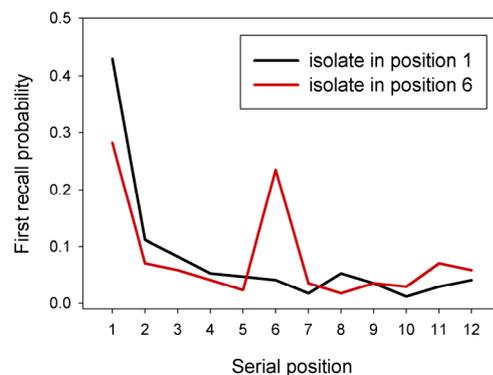
Figure 6.
Serial position curves of the data in Experiment 3. The isolate was presented either in position 1 or in position 6.

context model to which we now turn.

General Discussion

The three experiments provide evidence for the view that item-level novelty and context-level novelty enhance memory performance and contribute to primacy effects in free recall. As expected (Glenberg et al., 1983; Howard & Kahana, 1999; Murdock, 1972), the NICE context model produces recency effects. However, the important difference is that the probabilities of activating and de-activating contextual elements are made conditional on the novelty value between the previously and the newly presented items. As the first item is different from the information preceding it, its presentation coincides with an enhanced level of novelty, which in turn speeds up the change in episodic context. This additional change makes that first item more *locally distinctive* and more accessible during retrieval.

Although the begin- and end-points have been used in several models in memory retrieval (Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Davelaar et al., 2005; Henson, 1998; Metcalfe & Murdock, 1981; Neath, 1993; Shiffrin & Cook, 1977), none of these models provide a mechanistic explanation for the higher distinctiveness of the beginning and end of a list. The NICE context model provides such an explanation.

One of the assumptions is that novelty is measured as the difference in representation between the current and the previous item (plus background). This is consistent with single-store models that do not postulate a limited-capacity buffer. When a buffer is postulated, however, the novelty-signal could be computed from the current item and the buffer items. Such a change to the model does not invalidate the central thesis of this paper (i.e., Novelty can Induce Change in Episodic context) and may even provide a novel way to address the need for a limited-capacity buffer in understanding free recall performance.

In addition, it is assumed that stimulus representations contain semantic features of the to-be-remembered word and the features of the background picture on which the word is presented. Although this captures the data with a minimum set of assumptions, this operationalization of stimulus representations comes from the view that the background environment is also input and differs from the internal context representation. Despite the regular usage of context in computational models of memory, no clear understanding has been reached about what context is (and is not) (see for a recent review, Klein, Shiffrin & Criss, 2007).

Regarding novelty detection, it has been proposed that detection of novelty is related to increase in the release of acetylcholine, a neurotransmitter that originates from the basal forebrain and modulates several brain areas, including structures in the medial-temporal lobe. The mechanism by which the brain detects novelty is yet unclear, but there are several ways in which a novelty signal can be obtained in a neurocomputational model. For example, in ART models (Grossberg, 2012) an interplay between bottom-up input and top-down feedback leads to a novelty-signal when the top-down feedback does not resonate with the bottom-up input. In those models, the resulting novelty-signal is used to allow a new higher-level representation to become activated and get associated with the novel bottom-up input. A non-interactive (bottom-up and top-down) architecture is also possible when using units that adapt over time. This

adaptation (aka habituation, neural fatigue, synaptic depression) leads to the total activation to be an indicator of perceived stimulus novelty (Davelaar et al., 2011).

The NICE context model provides a solution to the question of how a context model can retrieve the first item of a preceding list. Jang and Huber (2008) replicated the so-called list-before-last experiments by Shiffrin (1970), showing that participants are able to skip the just encoded list and retrieve items from the preceding list. This provides a major challenge to all context models that are recency-biased. One solution is to assume that a hierarchy of contextual elements exists, such that different levels have faster or slower transition probabilities (cf. Davelaar & Usher, 2002; Glenberg et al., 1983). The modeler can then choose those elements that fluctuate in the same rhythm as the presentation of the trials. Another approach is to assume that each first item is more distinctive than the other items and thus can provide an anchor point during retrieval. The NICE context model produces memory traces for the first items of all lists that are locally distinctive. Thus retrieving the first item of the preceding list is easier than retrieving the middle item of the preceding lists, despite that the middle items are temporally more recent. It also naturally accounts for the observation that after retrieving items from the end of the list, the next item to be retrieved is most likely to be the very first list item (Laming, 1999).

Concluding Remarks

The current approach is not meant as an alternative to existing models, but instead introduces a mechanism by which other context models can capture primacy effects. For example, in TCM (Howard & Kahana, 2002), where contextual change is item-dependent, a similarity structure between the retrieved contexts of list items could be introduced. As the retrieved context vector of the first item differs more from that of the fixation stimulus than that of the second item, the change in the ongoing context will be larger for the first than the second item. It remains to be seen whether this modification to TCM would produce the desired results. Nevertheless, the basic concept will be the same, i.e., a source of contextual change is the novelty of an item at perceptual and conceptual levels.

Author Note

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1977). Recency re-examined. In S. Dornic (Ed.), *Attention and performance* (Vol. 6). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 176-189. doi:10.1016/S0022-5371(70)80048-2

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

Brodie, D. A., & Murdock, B. B. (1977). Effects of presentation time on nominal and functional serial position curves of free recall. *Journal of Verbal Learning and Verbal Behavior, 16,* 185-200. doi:10.1016/S0022-5371(77)80046-7

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114,* 539-576. doi:10.1037/0033-295X.114.3.539

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107,* 127-181. doi:10.1037/0033-295X.107.1.127

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551-581. doi:10.1037/0033-295X.106.3.551

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short term memory. *Journal of Verbal Learning and Verbal Behavior, 12,* 599-607. doi:10.1016/S0022-5371(73)80039-8

Davelaar, E. J., & Usher, M. (2002). An activation-based theory of immediate item memory. In J. A. Bullinaria, & W. Lowe (Eds.), *Connectionist models of cognition and perception: Proceedings of the seventh neural computation and psychology workshop*. Singapore City: World Scientific.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112,* 3-42. doi:10.1037/0033-295X.112.1.3

Davelaar, E. J., Tian, X., Weidemann, C. T., & Huber, D. E. (2011). A habituation account of change detection in same/different judgments. *Cognitive, Affective, & Behavioral Neuroscience, 11,* 608-626. doi:10.3758/s13415-011-0056-8

Debener, S., Herrman, C. S., Kranczioch, C., Gembris, D., & Engel, A. K. (2003). Top-down attentional processing enhances auditory evoked gamma band activity. *Neuroreport, 14,* 683-686. doi:10.1097/00001756-200304150-00005

Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review, 62,* 369-377. doi:10.1037/h0046888

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 231-255. doi:10.1037/0278-7393.9.2.231

Grossberg, S. (2012). Adaptive resonance theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks, 37,* 1-47. doi:10.1016/j.neunet.2012.09.017

Henson, R. H. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36,* 73-137. doi:10.1006/cogp.1998.0685

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. doi:10.1037/0278-7393.25.4.923

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

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review, 112,* 75-116. doi:10.1037/0033-295X.112.1.75

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27,* 1359-1366. doi:10.1037/0278-7393.27.6.1359

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. doi:10.1037/0278-7393.34.1.112

Kirchoff, B. A., Wagner, A. D., Maril, A., & Stern, C. E. (2000). Pre-

frontal-temporal circuitry for episodic encoding and subsequent memory. *Journal of Neuroscience, 20,* 6173-6180.

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

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34,* 419-426. doi:10.1080/002075999399774

Laming, D. (2006). Predicting free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1146-1163. doi:10.1037/0278-7393.32.5.1146

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

Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of single trial free recall. *Journal of Verbal Learning and Verbal Behavior, 20,* 161-189. doi:10.1016/S0022-5371(81)90365-0

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review, 67,* 16-31. doi:10.1037/h0042382

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Verbal Learning and Verbal Behavior, 64,* 482-488.

Murdock, B. B. (1972). Short-term memory. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 5). New York: Academic Press.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10,* 389-395. doi:10.1080/09658210244000216

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition, 21,* 689-698. doi:10.3758/BF03197199

Norman, K. A. & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems Approach. *Psychological Review, 110,* 611-646. doi:10.1037/0033-295X.110.4.611

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.

Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior, 14,* 623-629. doi:10.1016/S0022-5371(75)80049-1

Rundus, D. (1971). An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63-77. doi:10.1037/h0031185

Rundus, D., & Atkinson, R. C. (1970). Rehearsal procedures in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9,* 99-105. doi:10.1016/S0022-5371(70)80015-9

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20,* 11-21. doi:10.1136/jnnp.20.1.11

Sederberg, P. B., Gauthier, L. V., Terushkin, V., Miller, J. F., Barnathan, J. A., & Kahana, M. J. (2006). Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage, 32,* 1422-1431. doi:10.1016/j.neuroimage.2006.04.223

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

Shiffrin, R. M., & Cook, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior, 17,* 189-218. doi:10.1016/S0022-5371(78)90146-9

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

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory & Cognition, 30,* 885-892. doi:10.3758/BF03195774