A grand memory for forgetting:

Directed forgetting across contextual changes

Tracy L. Taylor[1]

& Jeff P. Hamm[2]

Keywords: Directed forgetting; intentional forgetting; item-method directed forgetting;

context effects

[1] Dalhousie University, Halifax, Nova Scotia, Canada

[2] The University of Auckland, Auckland, New Zealand

Correspondence should be addressed to:

Tracy Taylor
Department of Psychology and Neuroscience
Dalhousie University
1355 Oxford Street
Halifax, Nova Scotia, Canada
B3H 4R2
ttaylor2@dal.ca

Abstract

Using an item-method directed forgetting task, we presented homographic homophonic nouns embedded in sentences. At study, each sentence was followed by an instruction to remember or forget the embedded word. On a subsequent yes-no recognition test, each word was again embedded within a sentence. In Experiments 1, 2, and 4 we varied the embedding sentence at test so that it was identical to that at study, changed but retained the meaning of the studied word, or changed to alter the meaning of the studied word. Repeated context – whether the sentence and/or the word meaning – proved to be as useful a retrieval cue for TBF items as for TBR items. In Experiment 3, we demonstrated that physical repetition was insufficient to produce context effects for either TBR or TBF items. And, in Experiment 4, we determined that participants were equally accurate in reporting context repetition/change following the correct recognition of TBR and TBF items. When considered in light of the existing literature, our results suggest that when context can be dissociated from the study item, it is encoded in "one shot" and not vulnerable to subsequent efforts to limit unwanted encoding.

A grand memory for forgetting:

Directed forgetting across contextual changes

Intentional forgetting is an adaptive function motivated by the desire to retain a subset of items in long-term memory while discarding others that are deemed irrelevant or outdated (Anderson & Huddleston, 2012; Anderson, 2003; Bjork, 1989). In the laboratory, intentional forgetting can be studied using a directed forgetting paradigm (see MacLeod, 1998 for a review). The list-method version of this paradigm presents participants with a mid-list instruction to remember or forget all the items that they have already rehearsed and (presumably) encoded into long-term memory. Where a *directed forgetting effect* is operationalized as better memory performance for to-be-remembered (TBR) items than to-be-forgotten (TBF) items, a directed forgetting effect obtained using the list-method paradigm is attributed to inhibition (e.g., Geiselman, Bjork, & Fishman, 1983) and/or a mental context change (e.g., Sahakyan & Kelley, 2002) that reduces the likelihood that previously encoded TBF items will be retrieved. In contrast, the item-method version of the directed forgetting paradigm presents participants with study items one at a time, each followed with equal probability by an instruction to remember or forget. A directed forgetting effect in the item-method paradigm is attributed to selective rehearsal of TBR items to the relative exclusion of TBF items. This selective rehearsal operates at encoding to limit the commitment of unwanted TBF items to long-term memory. Thus, both the method and the underlying mechanism of directed forgetting differ depending on whether the intention to forget is formed *after* the TBF items have already been committed to memory (list method) or *before* the TBF items have been encoded (item method). The current study is concerned exclusively with intentional forgetting that operates to limit unwanted encoding – *before* the TBF items have been committed to memory. We are therefore interested exclusively in the item-method paradigm and its underpinnings.

Efficient control over encoding processes ensures that limited-capacity resources are not wasted rehearsing unwanted TBF items and are, instead, focused on rehearsing TBR items. To this end, an instruction to forget in the item-method paradigm engages frontal control mechanisms (e.g., Bastin, Feyers, Majerus, Balteau, Degueldre, Luxen, Maquet, Salmon, & Collette, 2012; Rizio & Dennis, 2013; van Hooff & Ford, 2011; Wylie, Foxe, & Taylor, 2008; Yang et al., 2012) in common with – albeit, not identical to (Fawcett & Taylor, 2010) – those used to prevent countermanded motor responses (e.g., Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Aron, Robbins, & Poldrack, 2004; Aron & Poldrack, 2006). Although this top-down control cannot prevent TBF item encoding entirely (Bancroft, Hockley, & Farquhar, 2013; Lee, Lee, & Tsai, 2007), it does limit further rehearsal (cf. Hourihan & Taylor, 2006) by initiating a withdrawal of attentional resources from the TBF item representation in working memory (Fawcett & Taylor, 2010; Taylor & Fawcett, 2011; Thompson, Hamm, & Taylor, 2014; Thompson & Taylor, 2015). Reflecting its episodic nature, this representation includes information about the location of the TBF item (Hourihan, Goldberg, & Taylor, 2007) as well as information about its perceptual attributes (Lee, Lee, & Fawcett, 2013). It also includes information about other items presented during the same encoding epoch. Accordingly, the withdrawal of attentional resources limits unwanted TBF item encoding as well as processing of non-study items that occur subsequently in close spatial/temporal proximity (Fawcett & Taylor, 2008, 2012; Lee & Hsu, 2012). Yet, in a test of context effects within an item-method directed forgetting task, Burgess, Hockley, and Hourihan (2017) argued that a forget instruction does not impact encoding of the context in which the TBF item is presented.

Context effects refer to improved episodic retrieval when study context is reinstated at test (e.g., Tulving & Thomson, 1973) – likely due to reactivation of the neural processing that had been associated with the studied item (e.g., Wing, Ritchey, & Cabeza, 2015). Burgess et al. (2017) presented an item-method directed forgetting task in which study words were superimposed on an otherwise irrelevant background picture. Words were later tested on a background that had been presented at study or else on a novel background that contained no picture. The results revealed a directed forgetting effect, with better recognition of TBR than TBF words. There was also evidence of context effects, with better recognition of words presented on a test background that was also presented at study compared to a novel background. There was, however, no significant interaction of memory instruction and context, with TBR and TBF recognition benefiting equally from the re-presentation of a background that had been presented at study. Burgess et al. (2017) interpreted this *null* interaction as support for Malmberg and Shiffrin's (2005) "one-shot" hypothesis that context is encoded in the first 1-2 seconds of stimulus presentation. When the memory instruction is delayed longer than this, context is already encoded – and, by implication, unaffected by processes initiated by the memory instruction. Consequently, repetition of the context at test is as effective at cueing recognition of TBF words as TBR words.

When describing context effects, a distinction is made between *context alpha* and *context beta* (Wickens, 1987). The classic example of context alpha comes from Godden and Baddeley (1975) who showed better memory for word lists that were studied and tested in the same environmental context – both on land or both under water – rather than in a different environmental context. Context alpha is independent of the stimulus but establishes a seemingly irrelevant background that nevertheless influences performance (cf. Baddeley, 1982). In contrast, context beta interacts with the item by providing meaning to an otherwise ambiguous stimulus. The Burgess et al. (2017) manipulation of the background image on which a study word was superimposed constituted a manipulation of context alpha. It is not clear whether their conclusions also hold for context beta. Indeed, where the embedding context interacts with the TBF item to disambiguate its meaning, it seems possible that the purported withdrawal of processing resources could render the embedding information as unavailable as the item itself.

To determine whether memory instructions interact with context beta, the current study presented participants with sentences one at a time. Each sentence served as the context for an embedded noun that was designated as the study item. Each noun was a homographic homophone whose meaning was disambiguated by the embedding sentence. Critically, these same nouns were also presented in a sentence context during a subsequent yes-no recognition test. The key manipulation was the relationship between the study and test sentence contexts. Nouns at test were a) embedded in the same sentence and thus shared the same meaning as at study (Same sentence, Same meaning: SsSm); b) embedded in a different sentence but nevertheless had the same meaning as at study (Different sentence, Same meaning: DsSm); or c) embedded in a different sentence that established a different meaning for the noun relative to study (Different sentence, Different meaning: DsDm). For example, if the designated study word was *coach* embedded in the sentence "Drinks and snacks were available in the third coach of the train", then at test participants might be presented with this same sentence again, in which case the sentence and the designated noun meaning remained unchanged (SsSm); an altered sentence in which the designated noun retained the same meaning, e.g., "The horse-drawn coach rattled noisily over the cobblestone drive" (DsSm); or, an altered sentence in which the designated noun assumed a different meaning, "Peter sometimes felt like he could use a life coach" or "The football coach was tough but the players liked her" (DsDm). Because each designated noun in our study had more than one possible interpretation, the embedding sentence served to constrain the interpretation and establish the intended meaning.

The critical question was whether the repetition of context – viz. repeated sentence (e.g., SsSm vs DsSm) and/or word meaning (e.g., DsSm vs DsDm) – was equally effective in cueing TBR and TBF item recognition. Certainly, there should be context effects for TBR items, with improved recognition due to contextual overlap between study and test. The prediction for the TBF items is less clear. On the one hand, if the attempt to limit unwanted encoding of an ambiguous TBF item also limits the encoding of its disambiguating context, overlap in study-test context should provide relatively little cueing of TBF item recognition. With benefits to TBR item recognition but not TBF item recognition, directed forgetting effects (defined as the difference in TBR and TBF item recognition) should therefore be larger when context repeats between study and test. On the other hand, if context beta – like context alpha – is encoded in "one-shot" within the first seconds of stimulus presentation, TBF items should benefit from contextual cueing as much as TBR items (e.g., Burgess et al., 2017). As a result, the magnitude of the directed forgetting effect should not vary as a function of context repetition across study and test.

EXPERIMENT 1

In Experiment 1, embedding sentences disambiguated the meanings of homographic homophones that served as study words in an item-method directed forgetting task. At test, these words were presented in one of three conditions: Same sentence-Same meaning (SsSm); Different sentence-Same meaning (DsSm); and, Different sentence-Different meaning (DsDm).

Note that there was no Same sentence-Different meaning (SsDm) condition as might be expected if our experimental manipulation of context reflected a factorial crossing of sentence identity and noun meaning. To understand this, it is critical to keep in mind that in all three of our context conditions – SsSm, DsSm, DsDm – the embedded noun repeated at study and test, even when the noun meaning changed. It is logically impossible for the same noun (e.g., *coach*) to appear in the same sentence at *both* study and test ("Drinks and snacks were available in the third ______ of the train") and yet have a different meaning in each instance. Thus, there was no possibility of including SsDm as one of the experimental context conditions in which the studied nouns repeated at test[1].

Method

[1] One might argue that even if it could not serve as an experimental context condition, the "missing" SsDm condition could nevertheless serve as a type of foil trial at recognition by replacing the studied noun with an unstudied noun. There are two problems with this. First, for the sentence to remain sensible generally requires inserting a synonym in place of the studied word – for example, the word *carriage* when *coach* refers to a conveyance and the word *instructor* when it does not. In this way, the sentence could remain the same at study and test (Ss); however, using a synonym for the studied word would not constitute a different meaning (Dm) for the embedded noun (as would be required to create a SsDm condition). Second, these synonyms would not necessarily be homographic homophones in each case. Participants could therefore reject these unstudied words on this basis alone, making them poor foils for testing recognition memory of studied words.

Participants. A total of 48 Dalhousie undergraduate students participated in exchange for course credit. All participants were tested individually in a session lasting less than 1hr. None reported having participated in other studies of directed forgetting and all participants self-reported comfort with the English language.

Stimuli and Apparatus. Stimuli consisted of a list of 96 nouns that had at least two alternate meanings in the noun form. These were randomized and distributed to one list of 24 TBR items, one list of 24 TBF items, and one list of 48 foil items. Study trials consisted of the intermixed presentation of the TBR and TBF items; test trials consisted of the intermixed presentation of the TBR, TBF, and foil items. For every noun, we constructed four embedding sentences. Two of these embedding sentences established one possible meaning of the noun (e.g., *coach* as a conveyance); two of these embedding sentences established an alternative meaning of the noun (e.g., *coach* as an instructor). The nouns and each of their four embedding sentences are shown in Appendix I.

To control for the frequency of each alternative noun meaning and for any vagaries of sentence construction, we carefully counterbalanced sentence presentation across all levels of word type (TBR, TBF, foil) and context (SsSm, DsSm, DsDm). This counterbalancing is represented in Table 1. This table shows the within-subject distribution of 96 nouns across TBR, TBF, and foil trials as a function of the embedding sentence (Sentence 1-4; see Appendix I) presented at study and at test.

Insert Table 1 about here

Using custom software written in Python, we also ensured that – across participants – individual words were presented equally often in all four of their embedding sentences and across levels of word type (TBR, TBF, foil) and context. To accomplish this, the software generated an initial random order of the 96 nouns, partitioning them into 8 consecutive lists of 12 nouns each. These eight 12-item sub-lists created through the initial noun randomization were numbered in sequence (1-8) and then placed in the following orders: 12345678, 23456781, 34567812, 45678123, 56781234, 67812345, 78123456, 81234567. Once the order was established, nouns on the first four sub-lists in the sequence were designated as study items and nouns on the last four sub-lists were designated as foils. Each one of the 4 study item sub-lists was mapped onto one of the 4 possible sentences (see Appendix I) – for example, Sentence 1 for the first study sub-list, Sentence 2 for the second study sub-list, etc. In each of the 4 study item sub-lists, 6 nouns were designated TBR and 6 were designated TBF.

Items within each of these TBR and TBF subsets were further defined by the sentence (see Appendix I) used to embed the noun at test. For example, for the 6 TBR and 6 TBF nouns embedded in Sentence 1 at study, 2 were likewise embedded in Sentence 1 at test (SsSm), 2 were embedded in Sentence 2 at test (DsSm), 1 was embedded in Sentence 3 at test (DsDm), and 1 was embedded in Sentence 4 at test (also DsDm). Determining which individual words within each 12-item sub-list served in each of these conditions was governed by crossing memory instruction (TBR, TBF) and context (SsSm, DsSm, DsDm). The 6 resulting permutations were applied to each of the 8 possible sub-list sequences, for a total of 48 unique combinations. Each of our 48 participants thus experienced a unique stimulus set.

For the foil items, each of the 4 sub-lists was assigned to a corresponding sentence number, such that the nouns on foil sub-list 1 were embedded in their corresponding Sentence 1 at test (again, please refer to Appendix I); nouns on foil sub-list 2 were embedded in their corresponding Sentence 2 at test; nouns on foil sub-list 3 were embedded in their corresponding Sentence 3 at test embedded; and, nouns on foil sub-list 4 were embedded in their corresponding Sentence 4 at test.

A separate piece of custom software, also written in Python, was used to present stimuli and record responses. This software was implemented on Apple iMacs equipped with a standard USB keyboard, and built-in 24-inch colour monitor.

Fixation crosshairs ("+") were drawn across the middle of the computer screen at the start of each trial and subtended approximately 1.5 degrees of visual angle, measured from a viewing distance of 57 cm. All text stimuli were printed in white colour on a black background in size-30 Helvetica font, written from left-to-right, and centered in the middle of the computer monitor. All study and test sentences were presented in sentence case, starting with a capitalized letter and ending with a period. The instruction to Remember was comprised of a string of 10 "R"s and the instruction to Forget was comprised of a string of 10 "F"s; these strings were likewise centered in the middle of the computer monitor.

Procedure. At the start of the experiment, participants were given a verbal overview that was reiterated by written instructions presented on the computer monitor. Participants were told that they would be presented with a series of sentences, one at a time, in which a word would be underlined. They were told that if the sentence was followed by a Remember instruction, they should commit the underlined word to memory for a subsequent memory test; if the sentence was followed by a Forget instruction, they were told that they could forget the underlined word. They were not forewarned that testing their compliance with these instructions required that their memory for the Forget words also be tested.

Figure 1 depicts the sequence of trial events. Each trial started with a 1,000-ms presentation of fixation crosshairs ("+") centered in the middle of the computer screen. The fixation stimulus was then replaced by a sentence in which the study word was embedded. This sentence was visible for 3,000 ms before underlining was used to identify the target word. We presented the sentence in advance of identifying the target noun to provide an opportunity for participants to read the sentence, even though they were not required to do so. We reasoned that reading would be relatively automatic (e.g., Stroop effect; see MacLeod, 1991) but that the inclination to read might be thwarted if immediate identification of the target noun allowed participants to narrow their focus of attention onto the noun exclusively (e.g., Laberge & Brown, 1986); if context were never processed, we would not be able to determine whether its repetition interacts with memory instruction. The sentence remained visible for a further 3,000 ms while the underlining remained visible. Underlining was used to identify the study word because it did not alter the width of the word to cause a horizontal displacement of the sentence in which it was embedded. Following removal of the sentence, the fixation stimulus reappeared for 500 ms before being replaced by a 500-ms presentation of the memory instruction.

Insert Figure 1 about here

After the presentation of all study trials, instructions for the test trials appeared alone on the computer monitor and were cleared by a press of the SPACE bar. These instructions informed participants that they would see sentences one at a time in which one word was underlined. They were to decide whether they recognised each underlined word from the earlier study trials – regardless of whether it had been designated TBR or TBF. Participants were instructed to press the "y" key for "yes" and the "n" key for "no".

As shown in Figure 1, each test trial began with the fixation crosshairs at center for 1,000 ms. The fixation stimulus was then replaced by a test sentence, above which participants were prompted with: "Do you recognise the underlined word from the earlier study trials of this experiment? (y/n):". A total of 96 test trials were presented, comprised of the 24 TBR words, 24 TBF words, and 48 foil words. For each participant, TBR and TBF were equally likely to be tested in the SsSm, DsDm, and DsDm conditions. Regardless of alterations of sentence context and word meaning, if a word was presented at test that had been presented at study, a "y" response on the recognition test was classified as a hit and an "n" response was classified as a miss. Responses to foil words were classified as a false alarm if the participant responded "y" and as a correct rejection if the participant responded "n".

Data Analysis. We performed all data collation and analyses using RStudio 1.1.383 (R Studio Team, 2016) running R 3.4.3 (R Core Team, 2017) and R packages plyr (Wickham, 2011), dplyr (Wickham & Francois, 2017), tidyr (Wickham, 2018), and stringr (Wickham, 2018). We used the R package ez 4.4-0 (Lawrence, 2016) to calculate descriptive statistics (ezStats) and within-subjects analyses of variance (ezANOVA), and to create plots (ezPlot) that were modified using ggplot2 (Wickham, 2009).

We used the output from ezANOVA to generate a Bayesian Information Criterion (BIC) approximation to Bayesian posterior probabilities according to the method described by Masson (2011). Unlike *null* hypothesis testing, which assesses the accuracy of the null hypothesis prediction, a Bayesian approach allows us to define the strength of support for or against the null hypothesis, given the evidence. This analytic approach is particularly useful given that the "one-shot" view of context encoding explicitly predicts a *null* interaction of memory instruction and context. However, recognising that not all readers will be familiar with this approach, we will also report the outcome of the analyses of variance (ANOVAs) that were used to generate the Bayesian approximations and provide generalised eta squared (*ges*) as a measure of effect size; nevertheless, all interpretations will be based on the Bayesian statistics.

We will use the convention *p*H1 to refer to the approximated posterior probability of the alternative (i.e., non-zero effect) given the data and *p*H0 to refer to the approximated posterior probability of a *null* effect given the data. Because the values of *p*H0 and *p*H1 sum to 1, we will report only the larger of the two values. When *p*H1 is the larger value (meaning there is *more* support for the alternative hypothesis of an effect than there is for the null hypothesis of no effect), we will describe our results in terms of evidence *for* an effect. When *p*H0 is the larger value (meaning there is *less* support for the alternative hypothesis of an effect than there is for the null hypothesis of no effect), we will describe our results in terms of evidence *against* an effect. In doing so, we will apply verbal descriptors using the conventions recommended by Raftery (1995): 0.50-0.75="weak", 0.75-0.95="positive", 0.95-0.99="strong", >0.99="very strong".

In an initial pass through the data, we calculated mean foil false alarm rates for each participant on the recognition test. If any participant's foil false alarm rate exceeded the mean of all participants by more than 2 standard deviations, the data contributed by that participant were excluded from all subsequent analyses.

Results

Our initial analysis of foil false alarm rates identified one dataset for exclusion due to a foil false alarm rate of 58%; subsequent analyses were performed using the remaining 47 datasets.

We analysed recognition hits as a function of memory instruction (TBR, TBF) and study-test context (SsSm, DsSm, DsDm). These data are shown in Figure 2. There was very strong evidence *for* a directed forgetting effect, with overall greater recognition of TBR words than TBF words, pH1>.99 [F(1,46)=55.47, p<.01, *ges*=.14]. There was also very strong evidence *for* an effect of context, pH1>.99 [F(2,92)=13.71, p<.01, *ges*=.06]. Pairwise comparisons revealed context effects due to both repeating the same sentence and repeating the same word meaning from study to test, albeit with more convincing evidence for the former. When word meaning remained unchanged from study to test, there was positive evidence *for* greater recognition hits when the sentence repeated (SsSm) rather than changed (DsSm), pH1=.94 [F(1,46)=10.01, p<.01, *ges*=.06]. And when the sentence changed from study to test, there was weak evidence *for* greater recognition hits when the word meaning repeated (DsSm) rather than changed (DsDm), pH1=.58 [F(1,46)=4.65, p<.04, *ges*=.02].

Critically, there was strong evidence *against* an interaction of memory instruction and context, pH0=.98 [F<1]. As shown in Figure 2, the magnitude of the directed forgetting effect (i.e., TBR hits – TBF hits) was 15% when the test words were embedded in the same sentence and retained the same meaning as at study (SsSm); 15% when the test words were embedded in a different sentence that otherwise conveyed the same meaning as at study (DsSm); and, 20% when the test words were embedded in a different sentence that conveyed a different meaning than at study (DsDm).

Insert Figure 2 about here

Discussion

The recognition results of Experiment 1 revealed an overall directed forgetting effect, suggesting that participants utilised the memory instructions to exert control over the contents of long-term memory. The results also revealed overall context effects, with better recognition performance when the study and test contexts matched rather than mismatched. These context effects reflected the influence of repeated sentence as well as repeated meaning. This was true despite the fact that participants were not explicitly required to encode the embedding sentences. Harkening back to the Stroop effect (MacLeod, 1991), participants clearly read these sentences when processing the embedded nouns and, in fact, were encouraged to do so by the delay that we imposed between presenting the sentence and designating the task-relevant noun.

The fact that the SsSm condition produced the best overall recognition performance compared to the DsSm and DsDm conditions would seem to suggest that this condition benefited from the combined effects of repeating both the embedding sentence and the word meaning. However, caution is needed in interpreting this result: The SsSm condition was the only one for which the embedding sentence was identical at study and test. In other conditions, including the foil trials, the sentence changed from study to test. As already detailed in our Introduction to Experiment 1, we were motivated against constructing foil trials on which the sentence stayed the same as at study but the embedded noun changed (i.e., the "missing" SsDm condition; see also Note 1). Nevertheless, this introduces the possibility that in the SsSm condition the sentence repetition itself biased participants to respond "yes" on the recognition task, without necessarily cueing study word retrieval.

That said, we do not view this possibility as condemnatory. This is because the repeated sentence could bias recognition responses in the SsSm condition only to the extent that sentence context was encoded at study and formed part of the item representation – thereby reinforcing, rather than undermining, a potential role of context in episodic memory representation. Indeed, even if we ignore the SsSm condition, and focus only on the DsSm condition for which meaning alone repeated, there was evidence for overall better recognition than in the DsDm condition for which no aspect of the study context repeated. This demonstrates that the repetition of meaning alone was sufficient to establish a context effect in overall recognition (see also Light & Carter-Sobell, 1970). Nevertheless, the concern that participants were especially aware of the sentence repetition in the SsSm condition will be addressed explicitly by the results of Experiment 4.

Regardless of how the main effect of context is ultimately interpreted, the core finding of Experiment 1 relates to the interaction of memory instruction and context. On this, our data were clear: TBF and TBF items benefitted equally from repeated context at study and test. This suggests that the TBF items that escape control at study are part of a rich contextual encoding that is likely content addressable (i.e., can be retrieved by accessing the representation of the context in which the study item was embedded, rather than only being retrieved by accessing the studied item itself). These unwanted memory traces thus do not appear to be spectres but fully-fledged memories that benefit from the recapitulation of context, just as if they had been intentionally encoded. This does not deny the possibility that some richness of detail is lost (Fawcett, Lawrence, Taylor, 2016) or that the TBF items are encoded more weakly than TBR items (Thompson, Fawcett, & Taylor, 2011). But there is clearly sufficient detail to support cueing of TBF items. This cueing may come from the repetition of the embedding context and/or the repetition of the disambiguated word meaning.

The results of Experiment 1 argue that context can cue retrieval of unwanted TBF memory traces as well as TBR memory traces. Before further discussing the implications of this finding, we thought it important to perform an independent replication of this interesting result. Experiment 2 thus exactly replicated the methods of Experiment 1 except that at test – and *only* at test – participants were alerted to the fact that the underlined words could be embedded in the same or different sentences relative to study and that their meaning could remain the same or change as a result. Our rationale for expanding the test instructions was that equivalent context effects for TBR and TBF items in Experiment 1 might have been due to participants using context rather than word strength to drive their recognition, despite our instruction to respond only to the underlined word. If so, context would have served as the *sole* retrieval cue for both TBR and TBF items rather than an *additional* retrieval cue.

EXPERIMENT 2

Experiment 2 was identical to Experiment 1, except for expanded test instructions.

Method

Participants. Initially, 48 Dalhousie undergraduate students participated in exchange for course credit. After completing their participation, two participants disclosed that they had volunteered in other studies of directed forgetting; their data were therefore excluded without ever being analysed and were replaced with datasets contributed by two additional recruits. Data were collected under similar conditions as described for Experiment 1.

Stimuli and Apparatus. The stimuli and apparatus were identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1, apart from expanded test instructions. The instructions that immediately preceded the recognition test alerted participants to the fact that sentences could remain the same or change from study to test and that word meanings could likewise stay the same or change. Participants were given examples of SsSm, DsSm, and DsDm conditions, using sentences constructed using the word *pipe*. The instructions emphasised that participants were to report recognition of all studied words, regardless of whether those words were designated TBR or TBF and regardless of whether the sentence and/or meaning changed. After participants depressed the SPACE bar, the test trials proceeded as described for Experiment 1. The only change was that the prompt underscored the need to assess recognition of the underlined word: "Do you recognise the underlined WORD from the earlier study trials of this experiment? (y/n):".

Results

The data from one participant were excluded from analysis due to a foil false alarm rate of 54%. The data from the remaining 47 participants were analysed in the same manner as described for Experiment 1.

We analysed recognition hits as a function of memory instruction (TBR, TBF) and study-test context (SsSm, DsSm, DsDm). These data are shown in Figure 3, along with the foil false alarm rate. There was very strong evidence *for* a directed forgetting effect, pH1>.99 [$F(1,46)=59.07$, $p<.01$, *ges*=.16]. There was also very strong evidence *for* context effects, pH1>.99 [$F(2,92)=17.88$, $p<.01$, *ges*=.05]. As shown in Figure 3, the effect of context reflects the fact that recognition hits were overall greatest when test words were presented in the same sentences and retained the same meaning as at study (M_{SsSm}=77%); of intermediate value when test words were presented in different sentences that nevertheless retained the same meaning as at study (M_{DsSm}=71%); and worst when test words were embedded in different sentences that conveyed different meanings than at study (M_{DsDm}=65%). As was the case for Experiment 1, pairwise comparisons revealed context effects due to both repeating the same sentence and repeating the same word meaning from study to test. When word meaning remained unchanged from study to test, there was positive evidence *for* greater recognition hits when the sentence repeated (SsSm) rather than changed (DsSm), pH1=.93 [$F(1,46)=9.64$, $p<.01$, *ges*=.03]; when the sentence changed from study to test, there was positive evidence *for* greater recognition hits when the word meaning repeated (DsSm) rather than changed (DsDm), pH1=.91 [$F(1,46)=8.97$, $p<.01$, *ges*=.03].

As in Experiment 1, there was strong evidence *against* an interaction of memory instruction and study-test context, pH0=.98 [$F<1$]. The magnitude of the directed forgetting effect (TBR hits – TBF hits) was 16% when the test words were presented in the same sentences that conveyed the same meaning as at study (SsSm); 17% when the test words were embedded in different sentences that nevertheless conveyed the same meaning as at study (DsSm); and, 21% when test words were embedded in different sentences that conveyed different meanings than at study (DsDm).

Insert Figure 3 about here

Discussion

As was the case for Experiment 1, the recognition results of Experiment 2 revealed an overall directed forgetting effect, with better recognition of TBR words than TBF words. There was also an overall context effect, with better recognition when the study and test contexts matched rather than mismatched. Indeed, recognition was better when the sentence itself repeated versus changed and also when the disambiguated word meaning repeated versus changed.

As was the case for Experiment 1, interpretation of the main effect of context should be tempered by the fact that participants might have approached recognition in the SsSm condition differently than in the other conditions for which there was no sentence repetition. We will again emphasize that this possibility only underscores and does not undermine the conclusion that context forms part of the episodic memory trace. Nevertheless, also as in Experiment 1, if we ignore the SsSm condition and focus only on the DsSm condition for which meaning alone repeated, there was better recognition than in the DsDm condition for which no aspect of the study context repeated. This again demonstrates that the repetition of meaning alone was sufficient to establish a context effect in overall recognition (see also Light & Carter-Sobell, 1970).

Regardless of the main effect of context in the omnibus analysis, the key question was whether context effects would interact with memory instruction in

Experiment 2. They did not. Despite more detailed test instructions that were designed to alert participants to potential changes in context and to focus them on word recognition, the conclusions of Experiment 2 echo those of Experiment 1. This was corroborated by adding Experiment as a between-subjects factor to the within-subjects analysis of memory instruction and study-test context. This analysis provided positive evidence *against* an effect of Experiment on overall recognition hit rates, pH0=.88 [F<1]; strong evidence *against* an interaction of Experiment with Memory Instruction, pH0=.90 [F<1]; very strong evidence *against* an interaction of Experiment with Study-Test Context, pH0>.99 [F<1]; and, most relevant, very strong evidence *against* a three-way interaction between Experiment, Memory Instruction, and Study-Test Context, pH0>.99 [F<1]. This very strong evidence against the three-way interaction argues that Experiments 1 and 2 produced a similar overall pattern of directed forgetting effects as a function of contextual overlap. Thus, we are confident that the results of *both* experiments provided material evidence against an interaction of memory instruction and context effects.

Together, Experiments 1 and 2 suggest that the repetition of sentence context serves as an effective cue to both TBR and TBF item recognition. It is not clear, however, whether it is the physical repetition of an embedding context that matters or whether it is the repetition of a disambiguating context *per se* that matters. To address this question, we replaced all of the characters in the embedding sentences with the letter 'x', while retaining the pattern of spacing, capitalization, and punctuation. In this way, the embedding context provided no guidance for interpreting the embedded noun yet could nevertheless physically repeat or change between study and test.

EXPERIMENT 3

Experiment 3 repeated the general methods of Experiments 1 and 2, except that the characters in the embedding sentences were changed to the letter 'x' in a case-sensitive manner.

Method

Participants. Initially, 48 Dalhousie undergraduate students participated in exchange for course credit. Two of these participants were accidentally run under the same counterbalancing manipulation; the second participant was thus excluded without any analysis of the contributed data and a new participant was recruited and run in the proper condition. A total of 48 data sets thus comprised the final sample.

Stimuli and Apparatus. All characters in each of the embedding sentences used in Experiments 1 and 2 were changed to 'x's on a case-sensitive basis, with no alteration in spacing or punctuation. Because the sentence could thus not establish the meaning of the embedded noun, the embedding context could only repeat or change from study to test; there was no corresponding repetition or change in word meaning. Nevertheless, to ease comparison with Experiments 1 and 2, we used the SsSm, DsSm, and DsDm conditions as dummy codes. This ensured that the trial composition was identical across experiments. We then renamed the conditions in Experiment 3 to reflect the nature of the repetition. The SsSm condition was renamed Ss to reflect that the embedding x-string mock sentences remained the same from study to test; the DsSm and DsDm conditions were renamed to Ds1 and Ds2, respectively, to reflect that in both data subsets the embedding mock x-string "sentences" were different between study and test.

The apparatus was identical to Experiment 2, except that the iMac computers were upgraded in the interim to 27" monitors.

Procedure. The procedure was identical to Experiment 1, with two exceptions. In the study and test instructions, the sentences were referred to as "mock sentences (consisting of strings of 'x's)". The test trial prompt was as described for Experiment 2.

Results

The data from four participants were excluded from analysis due to respective foil false alarm rates of 40%, 35%, 35%, and 31%. The data from the remaining 44 participants were analysed as for Experiment 1.

We analysed recognition hits as a function of memory instruction (TBR, TBF) and study-test context (Sm, Ds1, Ds2). These data are shown in Figure 4, along with the foil false alarm rate. There was very strong evidence *for* a directed forgetting effect, *p*H1>.99 [$F(1,43)=75.18$, $p<.01$, *ges*=.03]. Unlike in Experiments 1 and 2, however, there was strong evidence *against* context effects, *p*H0=.97 [$F<1$]. The hit rate averaged 64% when the embedding context stayed the same from study to test and a nearly identical 65% when the embedding context changed (i.e., M_{Ds1}=64%, M_{Ds2}=66%).

As in Experiments 1 and 2, there was strong evidence *against* an interaction of memory instruction and study-test context, *p*H0=.98 [$F<1$]. The magnitude of the directed forgetting effect (TBR hits – TBF hits) was 24% when the test words were presented in the same embedding context as at study (Ss) and averaged 23% when test words were presented in a different embedding context (M_{Ds1}=25%, M_{Ds2}=21%).

Insert Figure 4 about here

Discussion

In Experiment 3, the magnitude of the directed forgetting effect was ~23% whether the x-string repeated or changed. This is comparable to the magnitude of the directed forgetting effect that was revealed in the DsDm condition of Experiment 1 (20%) and Experiment 2 (21%). This suggests that participants in Experiment 3 were able to apply the memory instruction to the study words that were embedded in the nonsense x-strings but that there was no benefit to subsequent recognition due to repetition of those x-strings. This was true despite the fact that the pattern of capitalisation, spacing, and punctuation of the embedding x-strings matched those of the sentences used in Experiments 1-2.

Together, the results of Experiments 1-3 indicate that the repetition of a meaningful sentence context – not merely repetition of a physical context *per se* – provides a cue that is equally useful for the retrieval of TBR and TBF items. What is not clear is the extent to which participants are aware of the context repetition and whether their assessment of repetition/change is equally accurate for the recognition of TBR and TBF items. To address this, Experiment 4 repeated the methods of Experiment 2, except that after every item endorsement on the recognition task, participants were asked to indicate whether the embedding sentence was the same or different compared to study. The goal was to assess participants' accuracy in reporting the repetition/change of sentences that embedded TBR versus TBF words.

EXPERIMENT 4

Experiment 4 repeated the methods of Experiment 2 except that after every response that endorsed the recognition of a test word, participants were prompted to indicate whether the embedding context was the same or different relative to study. This query was intended to determine whether the subjective experience of recognising TBR and TBF items was accompanied by equally accurate assessments of the contextual change/repetition. The fact that context repetition is equally effective at cueing TBR and TBF item recognition does not necessarily imply that the representation of the embedding context is equally accessible to subjective awareness: Equivalent contextual cueing of TBR and TBF words might belie non-equivalent awareness of context repetition and change. If this were the case, it would suggest that the memory instruction does interact with contextual encoding, but at a different level of representation than needed to cue retrieval.

Method

Participants. A total of 48 Dalhousie undergraduate students participated in exchange for course credit. Data were collected under similar conditions as described for Experiment 1.

Stimuli and Apparatus. The stimuli were identical to Experiments 1 and 2; the apparatus was identical to Experiment 3.

Procedure. The procedure was identical to Experiment 2, except that after every "y" response on the recognition task the recognition prompt was replaced by a context prompt. This context prompt was printed in yellow to distinguish it from the preceding recognition prompt. While the test sentence (and underlined word) remained visible, participants were asked to indicate whether the word was presented in the same sentence as at study ("s") or in a different sentence ("d").

Results

The data from one participant were excluded from analysis due to a foil false alarm rate of 63%. Recognition and context judgment responses were analysed for the remaining 47 participants.

As for Experiments 1 and 2, we analysed recognition hits as a function of memory instruction (TBR, TBF) and study-test context (SsSm, DsSm, DsDm). These data are shown in Figure 5. There was very strong evidence *for* a directed forgetting effect, pH1>.99 [F(1,46)=24.67, p<.01, *ges*=.07]. There was also very strong evidence *for* context effects, pH1>.99 [F(2,92)=15.16, p<.01, *ges*=.06]. As shown in Figure 5, the effect of context reflects the fact that recognition hits were overall greatest when test words were presented in the same sentences and retained the same meaning as at study (M_{SsSm}=76%); of intermediate value when test words were presented in different sentences that nevertheless retained the same meaning as at study (M_{DsSm}=70%); and worst when test words were embedded in different sentences that conveyed different meanings than at study (M_{DsDm}=63%). Pairwise comparisons revealed context effects due to both repeating the same sentence and repeating the same word meaning from study to test. Even when word meaning remained unchanged from study to test, there was positive evidence *for* greater recognition hits when the sentence repeated (SsSm) rather than changed (DsSm), pH1=.87

[$F(1,46)=8.06$, $p<.01$, $ges=.03$]; when the sentence changed from study to test, there was positive evidence *for* greater recognition hits when the word meaning repeated (DsSm) rather than changed (DsDm), pH1=.85 [$F(1,46)=7.78$, $p<.01$, $ges=.03$].

As in Experiments 1 and 2, there was positive evidence *against* an interaction of memory instruction and study-test context, pH0=.94 [$F(2,92)=1.84$, $p>.16$, $ges<.01$]. The magnitude of the directed forgetting effect (TBR hits – TBF hits) was 9% when the test words were presented in the same sentences that conveyed the same meaning as at study (SsSm); 11% when the test words were embedded in different sentences that nevertheless conveyed the same meaning as at study (DsSm); and, 15% when test words were embedded in different sentences that conveyed different meanings than at study (DsDm).

Insert Figure 5 about here

Following "yes" responses on the recognition task, participants were asked to identify whether the context was the "same" or "different" relative to study. Because foil words were presented only at test and never at study, it was impossible to assess accuracy for responses that indicated the context had remained the "same" (M=20%) or was "different" (M=80%), since neither was true[2]. The mean accuracy for making the context determination was therefore only calculated following hits (i.e., "yes" recognition responses to studied TBR and TBF words). These data are shown in Table 2 and were analysed as a function of memory instruction (TBR, TBF) and study-test context (SsSm, DsSm, DsDm). With a mean accuracy of 75% following TBR item recognition and 77% following TBF item recognition, this analysis provided positive evidence *against* an effect of memory instruction, pH0=.81 [$F<1$]. There was, however,

[2] Two participants produced no false alarms to foils and so were not queried for foil context. These "same" and "different" context judgments to foil words therefore reflect averages calculated over 45 participants.

positive evidence *for* an effect of context, *p*H1=.93 [$F(2,92)=7.56$, $p<.01$, *ges*=.07]. After endorsing an item as having been studied, participants correctly identified the relationship between the study and test contexts on 67% of SsSm trials, compared to 78% of DsSm trials and 82% of DsDm trials. Pairwise comparisons revealed that when word meaning remained the same from study to test, there was weak evidence *for* poorer accuracy in reporting a repeated sentence (SsSm) versus a changed sentence (DsSm), *p*H1=.68 [$F(1,46)=5.56$, $p<.02$, *ges*=.07]. In contrast, when the sentence changed from study to test, there was weak evidence *against* a difference in accuracy when the word meaning repeated (DsSm) rather than changed (DsDm), *p*H0=.65 [$F(1,46)=2.59$, $p>.11$, *ges*=.01]. In other words, participants were overall more accurate to report a changed context than one that stayed the same between study and test. Critically, this tendency was unaffected by whether the recognised word was a TBR or a TBF item, as indicated by positive evidence *against* an interaction of memory instruction and context change, *p*H0=0.94 [$F(2,92)=1.72$, $p>.18$, *ges*<.01].

Insert Table 2 about here

Discussion

The recognition data of Experiments 1, 2, and 4 provide three independent replications that converge on the conclusion that memory instructions and context effects do not interact. The repetition of a disambiguating sentence from study to test facilitates greater recognition of the embedded item and does so to the same extent whether that item is designated TBR or TBF. When the word meaning repeats, recognition is improved by repeating versus changing the embedding sentence; when the embedding sentence changes, recognition is improved by repeating versus changing the embedded word meaning.

A visual comparison of the DsSm and DsDm conditions in Experiments 1, 2, and 4 (all of which showed context effects) seems to suggest that the effect of retaining the same word meaning in a changed sentence has a larger effect on TBF words than TBR words. To test this, we analysed recognition hits as a function of memory instruction (TBR, TBF), study-test context (DsSm, DsDm), and experiment (1, 2, 4). Underscoring the directed forgetting and context effects that were observed in each individual experiment, over all three experiments there was very strong support *for* an effect of memory instruction, *p*H1>.99 [*F*(1,138)=128.80, *p*<.01, *ges*=.13] and *for* an effect of repeating versus changing the meaning of the word in a changed sentence, *p*H1>.99 [*F*(1,138)=20.64, *p*<.01, *ges*=.02]. There was, however, weak evidence *against* an interaction of memory instruction and context, *p*H0=.75 [*F*(1,138)=2.79, *p*>.09, *ges*<.01]. This indicates that repeating the meaning in a changed sentence did not, in fact, affect TBF words more than TBR words. In this analysis, there was also strong evidence *against* an effect of experiment, *p*H0=.99 [*F*<1]; strong evidence *against* an interaction of memory instruction with experiment, *p*H0=.97 [*F*(2,138)=1.38, *p*>.25, *ges*<.01]; strong evidence *against* an interaction of context condition with experiment, *p*H0=.99 [*F*<1]; and, very strong evidence *against* a three-way interaction of memory instruction, context, and experiment, *p*H0>.99 [*F*<1].

Interestingly, when Experiment 4 followed up "yes" recognition responses by asking participants to report whether the sentence context was the "same" or "different" compared to study, responses were more accurate to report a change ("different") than a repetition ("same"). That said, participants were as likely to respond "different" after reporting a false recognition of unstudied foil words

(*M*=80%) as they were to respond "different" after reporting accurate recognition of studied words in the DsSm (*M*=78%) and DsDm (*M*=82%) conditions. Yet, when the context remained unchanged, participants detected this repetition with 67% accuracy, reporting "different" on only 33% of trials. This suggests that participants have some awareness of context repetition and could potentially use this information heuristically to drive increased recognition in the SsSm condition. Critical for present purposes, however, is the fact that context – however it is ultimately utilised – is represented similarly for TBR and TBF words: There was no evidence that the accurate report of context was affected by memory instruction. When participants correctly recognised a TBF item as having been presented at study, the embedding context was as available for explicit report (and potential use as a recognition heuristic) as it was for a recognised TBR item.

General Discussion

Robert Louis Stevenson claimed that he had a "grand memory for forgetting". Our results underscore the fact that most of us do. Our participants were able to flexibly control the contents of their long-term memory by committing items that received a remember instruction and by preventing the commitment of items that received a forget instruction. The adaptability of the memorial system for top-down control over the selection of relevant over irrelevant inputs is evident (Anderson, 2003). But the fact that we can selectively remember and forget individual words that are disambiguated by an embedding sentence context is quite remarkable.

The recognition results of Experiments 1, 2, and 4 established that repetition of a disambiguating encoding context cues subsequent recognition of both TBR and

TBF words, and that it does so approximately equally. Experiment 3 demonstrated that this was not due to mere physical repetition of an embedding context from study to test: Recognition of TBR and TBF words was unaffected by whether an embedding mock sentence of x-strings repeated or changed. Finally, Experiment 4 determined that when participants correctly recognised a studied word they were more accurate to report when the embedding context had changed rather than remained the same; this was true whether the recognised word was designated TBR or TBF.

These findings corroborate and extend those reported by Burgess et al. (2017). Whereas Burgess et al. failed to reject a *null* interaction of memory instruction and context alpha, our results provide evidence *against* an interaction of memory instruction and context beta. Concordant with their conclusion in favour of the "one shot" hypothesis, we presume that context beta – like context alpha – is encoded in the first moments following presentation. When a memory instruction follows, it is applied to the embedded word without influencing the embedding context. Consequently, there is a directed forgetting effect with better subsequent recognition of TBR words than TBF words. There is also an effect of context, with better recognition due to repetition of the embedding sentence and/or disambiguated word meaning. But there is no interaction of memory instruction and sentence context. This suggests that processing the embedding sentence context is neither truncated by the presentation of a forget instruction, nor elaborated further by the presentation of a remember instruction. The fact that this seems to be the case for both context alpha and context beta helps establish important limits for what information is and is not intentionally forgotten.

Interestingly, when context alpha is manipulated in an item-method directed forgetting task by presenting test words in the same or a different spatial location than they occupied during study, TBF items show a particular benefit in recognition compared to TBR items (Hourihan et al., 2007). This is consistent with the fact that short-term verbal memory representations include obligatory encoding and storage of context alpha (e.g., Postle, 2003). And it suggests that spatial location – despite being task-irrelevant – is represented along with the word at study so that it ultimately becomes part of the long-term episodic representation that is formed for that word.

On the surface, it seems a bit puzzling that a repeated spatial context is more effective at retrieving TBF than TBR items, whereas an unrelated background image context (Burgess et al., 2017) and a disambiguating sentence context are both equally effective at cueing retrieval of TBR and TBF items. On the one hand, this could be used to argue against the Burgess et al. (2017) interpretation of a *null* interaction of memory instruction and context alpha. On the other hand, it is reasonable to presume that the interaction of memory instruction and spatial context occurs because space is an indespensible attribute of vision (Kubovy, 1988): The visual representation of a peripherally presented study word is inexorably tied to the location of its presentation. In contrast, the background image used to establish context alpha in the Burgess et al. (2017) study might not have formed a cohesive representation with the foreground words that served as TBR and TBF items. Accordingly, relative immunity of contextual encoding from the impact of subsequent memory instructions might depend critically on the ability to dissociate the study item from its context – whether perceptually or conceptually.

Consider work by Fawcett et al. (2016). They presented visual roulettes (e.g., 'spirographs') at study, each of which was drawn in a unique colour. At test, participants were presented with white roulettes. For each test roulette, participants had to determine whether the item had been presented at study or not and then had to select its remembered colour from a continuous hue-chroma-luminance colour wheel – even though colour had not been explicitly task-relevant at study. Revealing a typical directed forgetting effect, participants recognised fewer TBF items than TBR items. But even when participants correctly recognised TBF items as having been presented at study, their colour selection was less accurate than for recognised TBR items. This demonstrated that the forget instruction impacted not only the probability of subsequent recognition, but also the fidelity of those memory traces that were successfully recognised.

In Fawcett et al.'s (2016) study, the coloured lines that comprised the roulettes were task-irrelevant at study but nevertheless integral to defining the study items: The roulettes did not exist apart from the spatial arrangement of coloured lines. In contrast, the sentences in the current experiments provided context and meaning to the embedded study words but did not share the same spatial location as those words and were not critical for identifying those words (even if they were critical for disambiguating their intended *meaning*). The fact that we did not observe an interaction of memory instruction with context suggests that a forget instruction might lead to the loss of perceptual details associated with the study item (as for Fawcett et al.'s roulettes), without necessarily affecting the encoding of conceptual information that establishes the *meaning* of the study item.

Supporting this possibility, Fawcett, Taylor, and Nadel (2013a, 2013b) extended the typical item-method paradigm by presenting visual vignettes with both diegetic and non-diegetic sound removed. These visual vignettes depicted a continuous event sequence (e.g., baking cookies) and participants were instructed to remember and forget segments of these vignettes. The memory instruction was communicated during a brief interruption of the ongoing event sequence (Fawcett et al., 2013a) or by an alteration in the colour of a visual frame that surrounded the video clip (Fawcett et al., 2013b). In both cases, participants showed better overall memory for TBR segments than for TBF segments. However, when Fawcett et al. (2013a, 2013b) manipulated the test questions to assess general or specific information, participants showed a directed forgetting effect for the specific representations (e.g., *The recipe called for 2 tablespoons of cornstarch*) but not for the general representations (e.g., *The recipe called for cornstarch*). In other words, what participants seemed to intentionally forget were the perceptual details of the video sequences rather than their conceptual understanding of those sequences: They understood that cornstarch was added even if they could not remember seeing how much was added.

It thus seems that forget instructions impact memory for details of visually presented event sequences (e.g., Fawcett et al., 2013a, 2013b), enable effective contextual cueing by the repetition of visuo-spatial location (e.g., Hourihan et al., 2007), and decrease memory for visual perceptual details (e.g., Fawcett et al., 2016), without necessarily reducing contextual cueing by either context alpha or context beta. This accords with a distinction made by Ecker, Zimmer, and Groh-Bordin (2007) between intrinsic and extrinsic feature binding in episodic memory. They argue that intrinsic object features (e.g., colour, spatial location) are bound into a unitary *object token*, the components of which are activated during retrieval. This implies that when a weak memory trace is activated (such as for TBF words), the intrinsic features (e.g., spatial location) are potentially available to drive recognition of the object token (cf. Hourihan et al., 2007). Success in doing so, of course, might very well depend on the extent to which the fidelity of the intrinsic feature representation (e.g., colour) has been compromised by the instruction to forget (e.g., Fawcett et al. 2016). In contrast, extrinsic features that comprise the encoding context are bound into an *episodic token* that is activated during retrieval through voluntary access. Our Experiment 4 results suggest that TBF words are less likely to be encoded than TBR words (i.e., a directed forgetting effect) but that there is unimpeded voluntary access to the contextual information contained within those episodic tokens that are successfully retrieved (i.e., equally accurate context change recognition for TBR and TBF words).

The fact that perceptual representations are more vulnerable to an intention to forget is consistent with the suggestion that forget instructions enacted at encoding initiate a withdrawal of visuo-spatial attention (Taylor & Fawcett, 2011; Taylor, 2005; Thompson et al., 2014; Thompson & Taylor, 2015). This truncates further TBF item processing (Taylor, in press) and decreases the processing of other information presented subsequent in close spatial and/or temporal proximity (Fawcett & Taylor, 2008, 2012; Lee & Hsu, 2012). But it does not operate retroactively to impact the representation of contextual details that were encoded prior to the memory instruction. Instead, these contextual details remain content-addressable and can cue subsequent recognition, despite an overall lower probability of successfully retrieving TBF item representations versus TBR item representations. Of course, as demonstrated by the results of Experiment 3, this is true only to the extent that the contextual information is inherently meaningful.

The fact that a repetition of sentence context between study and test aided subsequent recognition of both TBF and TBR items underscores the distinction between intentional forgetting and unintentional forgetting. One cause of unintentional forgetting is a lack of appropriate cues to retrieval (Tulving, 1974). When the retrieval context is different from the encoding context, a performance deficit may emerge due to an inability to retrieve a successfully encoded memory.

Overcoming this deficit may require reinstating the context either physically (see Smith & Vela, 2001 for a review) – as when one returns to a spatial location to try to trigger a memory for events encoded there – or mentally (Smith, 1979, 1984). If intentional forgetting in our task shared mechanisms in common with unintentional forgetting, we would have expected to observe an interaction of memory instruction and context change. Because this did not occur – and no matter the reason – it emphasises that item-method directed forgetting likely depends on encoding-based strategies to the exclusion of retrieval-based strategies (see also MacLeod, 1975; Taylor, Cutmore, & Pries, 2018) – despite the fact that neural signatures (Nowicka, Marchewka, Jednoróg, Tacikowski, & Brechmann, 2011; Paz-Caballero & Menor, 1999; Ullsperger, Mecklinger, & Müller, 2000; van Hooff & Ford, 2011) and multinomial modelling efforts (Rummel, Marevic, & Kuhlmann, 2016) suggest otherwise.

In short, our results demonstrate that context is encoded not only for items that we wish to remember but also for items that we fail to forget. When the encoded context is inherently meaningful, it enables content-addressable retrieval, such that repeating the study context at test helps cue recognition regardless of the memory intention instantiated at encoding. When context cues the retrieval of items we wish to remember, we enjoy an unexpected benefit from the processing of the embedding context. However, when context cues the retrieval of items that we intend to forget – and might even have forgotten were it not for the repetition of context – it undermines our intentions. In so doing, context offers another avenue for the retrieval of those items that escape selection at encoding, and reinforces retrieval of ostensibly weak memory traces.

Acknowledgements

We would like to thank Tom Howells for creating the sentences used in this experiment and for contributing to the data collection; Kate Thompson, Laura Cutmore, and Colin McCormick for help with data collection; and, Carl Helmick for writing the custom software. Thanks also to our participants for contributing their time and effort to this research project and to Dr. Maciej Hanczakowski and an anonymous reviewer for feedback on this work. This research was supported by a Discovery grant awarded to TLT from the Natural Sciences and Engineering Research Council of Canada.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*(4), 415-445.

Anderson, M. C., & Huddleston, E. (2012). Towards a cognitive and neurobiological model of motivated forgetting. In *True and false recovered memories* (pp. 53-120). Springer, New York, NY.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*(2), 115–116.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *The Journal of Neuroscience*, *26*(9), 2424–2433.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Baddeley, A. D. (1982). Domains of recollection. *Psychological Review*, *89*, 708-729.

Bancroft, T. D., Hockley, W. E., & Farquhar, R. (2013). The longer we have to forget the more we remember: The ironic effect of postcue duration in item-based directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 691-699.

Bastin, C., Feyers, D., Majerus, S., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., Salmon, E., & Collette, F. (2012). The neural substrates of memory suppression: a FMRI exploration of directed forgetting. *PloS One*, *7*(1), e29905.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Erlbaum.

Burgess, N., Hockley, W. E., & Hourihan, K. L. (2017). The effects of context in item-based directed forgetting: Evidence for "one-shot" context storage. *Memory & Cognition*, *45*(5), 745-754.

Ecker, U. K. H., Zimmer, H. D., & Groh-Bordin, C. (2007). Color and context: An ERP study on intrinsic and extrinsic feature binding in episodic memory. *Memory & Cognition*, *35*(6), 1483-1501.

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, *36*(6), 1168–1181.

Fawcett, J. M., & Taylor, T. L. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not with stop-signal inhibition. *Memory & Cognition*, *38*(6), 797–808.

Fawcett, J. M., & Taylor, T. L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica*, *139*(1), 84–90.

Fawcett, J. M., Taylor, T. L., & Nadel, L. (2013a). Event-method directed forgetting: Forgetting a video segment is more effortful than remembering it. *Acta Psychologica*, *144*(2), 332–343.

Fawcett, J. M., Taylor, T. L., & Nadel, L. (2013b). Intentional forgetting diminishes memory for continuous events. *Memory*, *21*(6), 675–694.

Fawcett, J. M., Lawrence, M. A., & Taylor, T. L. (2016). The representational consequences of intentional forgetting: Impairments to both the probability and fidelity of long-term memory. *Journal of Experimental Psychology: General*, *145*(1), 56-81.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: a link with posthypnotic amnesia. *Journal of Experimental Psychology: General*, *112*(1), 58-72.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent Memory in Two Natural Environments: On Land and Underwater. *British Journal of Psychology*, *66*(3), 325–331.

Hourihan, K. L., Goldberg, S., & Taylor, T. L. (2007). The role of spatial location in remembering and forgetting peripheral words. *Canadian Journal of Experimental Psychology*, *61*(2), 91–101.

Hourihan, K. L., & Taylor, T. L. (2006). Cease remembering: control processes in directed forgetting. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(6), 1354–1365.

Kubovy, M. (1988). Should we resist the seductiveness of the space: time:: vision: audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 318-320.

Laberge, D., & Brown, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. *Perception & Psychophysics*, *40*, 188-200.

Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. R package version 4.4-0. http://CRAN.R-project.org/package=ez

Lee, Y., & Hsu, Y. (2012). How do we forget negative events? The role of attentional, cognitive, and metacognitive control. *Cognition & Emotion*, *27*(3), 1–15.

Lee, Y. S., Lee, H. M., & Fawcett, J. M. (2013). Intentional forgetting reduces color-naming interference: Evidence from item-method directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 220-236.

Lee, Y., Lee, H., & Tsai, S. (2007). Effects of post-cue interval on intentional forgetting. *British Journal of Psychology*, *98*(2), 257–272.

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 1-11.

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(3), 271-279.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional* Forget*ting: Interdisciplinary approaches* (pp. 1- 57). Mahwah, NJ: Erlbaum.

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

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679-690.

Nowicka, A., Marchewka, A., Jednoróg, K., Tacikowski, P., & Brechmann, A. (2011). Successfully overcoming the inhibitory impact of the "forget" instruction: A voxel-based morphometric study of directed forgetting. *Psychophysiology*, *46*, 1108-1112.

Paz-Caballero, M., & Menor, J. (1999). ERP correlates of directed forgetting effects in direct and indirect memory tests. *European Journal of Cognitive Psychology*, *11*(2), 239-260.

Postle, B. R. (2003). Context in verbal short-term memory. *Memory & Cognition*, *31*(8), 1198-1207.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111-164.

Rizio, A. A., & Dennis, N. A. (2013). The neural correlates of cognitive control: successful remembering and intentional forgetting. *Journal of Cognitive Neuroscience*, *25*(2), 297-312.

Rummel, J., Marevic, I., & Kuhlmann, B. G. (2016). Investigating storage and retrieval processes of directed forgetting: A model-based approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(10), 1526-1543.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064-1072.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning & Memory*, *5*(5), 460–471.

Smith, S. M. (1984). A comparison of two techniques for reducing context-dependent forgetting. *Memory & Cognition*, *12*(5), 477–482.

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

Taylor, T. L. (2005). Inhibition of return following instructions to remember and forget. *The Quarterly Journal of Experimental Psychology*, *58*(4), 613–629.

Taylor, T. L. (in press). Remember to blink: Reduced attentional blink following instructions to forget. *Attention, Perception, & Psychophysics*.

Taylor, T. L., Cutmore, L., & Pries, L. (2018). Item-method directed forgetting: Effects at retrieval? *Acta Psychologica*, *183*, 116-123.

Taylor, T. L., & Fawcett, J. M. (2011). Larger IOR effects following forget than following remember instructions depend on exogenous attentional withdrawal and target localization. *Attention, Perception & Psychophysics*, *73*(6), 1790–814.

Thompson, K. M., Fawcett, J. M., & Taylor, T. L. (2011). Tag, you're it: tagging as an alternative to yes/no recognition in item method directed forgetting. *Acta Psychologica*, *138*(1), 171–175.

Thompson, K. M., Hamm, J. P., & Taylor, T. L. (2014). Effects of memory instruction on attention and information processing: Further investigation of inhibition of return in item-method directed forgetting. *Attention, Perception & Psychophysics*, *76*(2), 322–334.

Thompson, K. M., & Taylor, T. L. (2015). Memory instruction interacts with both visual and motoric inhibition of return. *Attention, Perception, & Psychophysics*, *77*, 804–818.

Tulving, E. (1974). Cue-dependent forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *American Scientist*, *62*(1), 74-82.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373.

Ullsperger, M., Mecklinger, A., & Müller, U. (2000). An electrophysiological test of directed forgetting: the role of retrieval inhibition. *Journal of Cognitive Neuroscience*, *12*(6), 924–940.

Van Hooff, J. C., & Ford, R. M. (2011). Remember to forget: ERP evidence for inhibition in an item-method directed forgetting paradigm. *Brain Research*, *1392*, 80–92.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer New York. http://had.co.nz/ggplot2/book

Wickham, H. (2011). The split-apply-combine strategy for data analysis. Journal of Statistical Software, 40(1), 1-29. http://www.jstatsoft.org/v40/i01/.

Wickham, H. (2018). stringr: Make it easier to work with strings. R package version 1.3.0. https://CRAN.R-project.org/package=stringr

Wickham, H. (2018). tidyr: Easily Tidy Data with spread() and gather() Functions. R package version 0.8.0. https://CRAN.R-project.org/package=tidyr

Wickham, H. & Francois, R. (2017). dplyr: A grammar of data manipulation. R package version 0.7.4. https://CRAN.R-project.org/package=dplyr

Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *Journal of Cognitive Neuroscience*, *27*(4), 679-691.

Wylie, G. R., Foxe, J. J., & Taylor, T. L. (2008). Forgetting as an active process: an FMRI investigation of item-method-directed forgetting. *Cerebral Cortex*, *18*(3), 670–682.

Yang, W., Liu, P., Xiao, X., Li, X., Zeng, C., Qiu, J., & Zhang, Q. (2012). Different neural substrates underlying directed forgetting for negative and neutral images: an event-related potential study. *Brain Research*, *1441*, 53–63.

Table 1. Within-participant distribution of 96 nouns across trials. The number in each cell represents the number of nouns corresponding to each combination of word type, study sentence, and test sentence. The sentences are numbered 1-4 and correspond with the four sentences created for each of the nouns, as shown in Appendix I.
TBR=to-be-remembered words, TBF=to-be-forgotten words, SsSm=Same sentence Same meaning, DsSm = Different sentence Same meaning, DsDm = Different sentence different meaning

		Test Sentence			
		Same Meaning		Different Meaning	
Word Type	Study Sentence	Sentence 1	Sentence 2	Sentence 3	Sentence 4
TBR	Sentence 1	2 (SsSm)	2 (DsSm)	1 (DsDm)	1 (DsDm)
	Sentence 2	2 (DsSm)	2 (SsSm)	1 (DsDm)	1 (DsDm)
	Sentence 3	1 (DsDm)	1 (DsDm)	2 (SsSm)	2 (DsSm)
	Sentence 4	1 (DsDm)	1 (DsDm)	2 (DsSm)	2 (SsSm)
TBF	Sentence 1	2 (SsSm)	2 (DsSm)	1 (DsDm)	1 (DsDm)
	Sentence 2	2 (DsSm)	2 (SsSm)	1 (DsDm)	1 (DsDm)
	Sentence 3	1 (DsDm)	1 (DsDm)	2 (SsSm)	2 (DsSm)
	Sentence 4	1 (DsDm)	1 (DsDm)	2 (DsSm)	2 (SsSm)
Foil		12	12	12	12

Table 2. Mean percentage "different" responses to the test context in Experiment 4; standard deviations shown in parentheses. These represent incorrect responses for TBR and TBF items in the SsSm context condition but correct responses in the DsDm and DsDm conditions. The mean percentage "different" responses are also shown for unstudied Foil words. *TBR=to-be-remembered words, TBF=to-be-forgotten words, SsSm=Same sentence Same meaning, DsSm = Different sentence Same meaning, DsDm = Different sentence different meaning*

		Context		
	Word Type	**SsSm**	**DsDm**	**DsDm**
Experiment 4	TBR	32 (24)	80 (16)	76 (19)
	TBF	33 (27)	63 (24)	55 (23)
	Foil		80 (20)	

"Different" responses were incorrect in the SsSm condition for which the sentence context actually remained the same: The corresponding accuracies are calculated by taking the difference from 100 and are 68% for TBR words and 67% for TBF words. "Different" responses were neither correct nor incorrect for unstudied Foil words.

Figure Captions

Figure 1

Schematic representation of the study and recognition trials. Stimuli are not drawn to scale and colours are inverted to show white on black instead of black on white as presented in the experiment.

Figure 2

Recognition hits (%) on the yes-no recognition test in Experiment 1, as a function of memory instruction (TBR, TBF) and study-test context (SsSm=Same meaning, same sentence; DsSm = Same meaning, different sentence; DsDm = Different meaning, different sentence). The dashed horizontal line represents the overall false alarm rate to foils. Error bars depict Fisher's Least Significant Difference for the plotted effect.

Figure 3

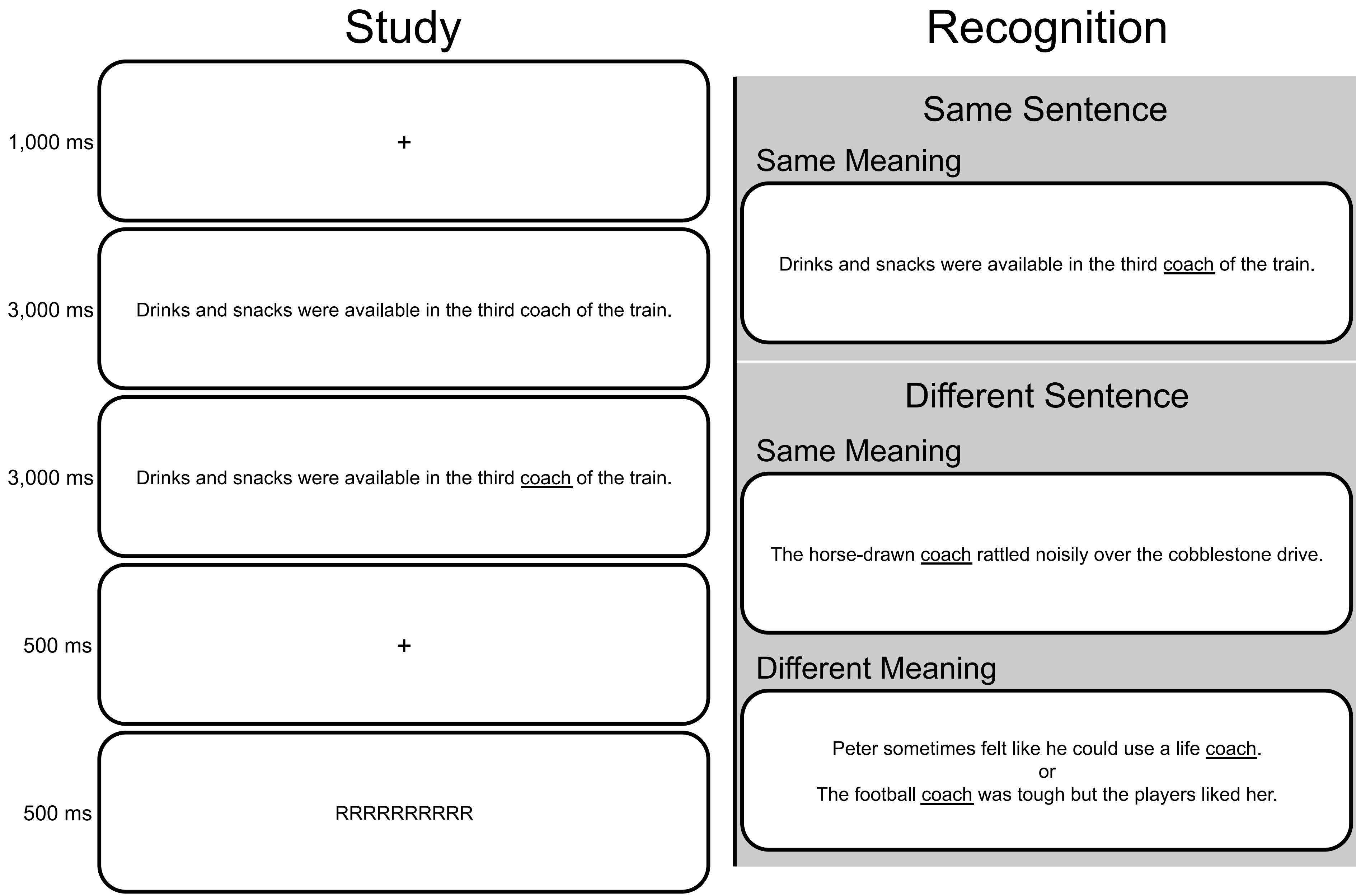
Recognition hits (%) on the yes-no recognition test in Experiment 2, as a function of memory instruction (TBR, TBF) and study-test context (SsSm=Same meaning, same sentence; DsSm = Same meaning, different sentence; DsDm = Different meaning, different sentence). The dashed horizontal line represents the overall false alarm rate to foils. Error bars depict Fisher's Least Significant Difference for the plotted effect.

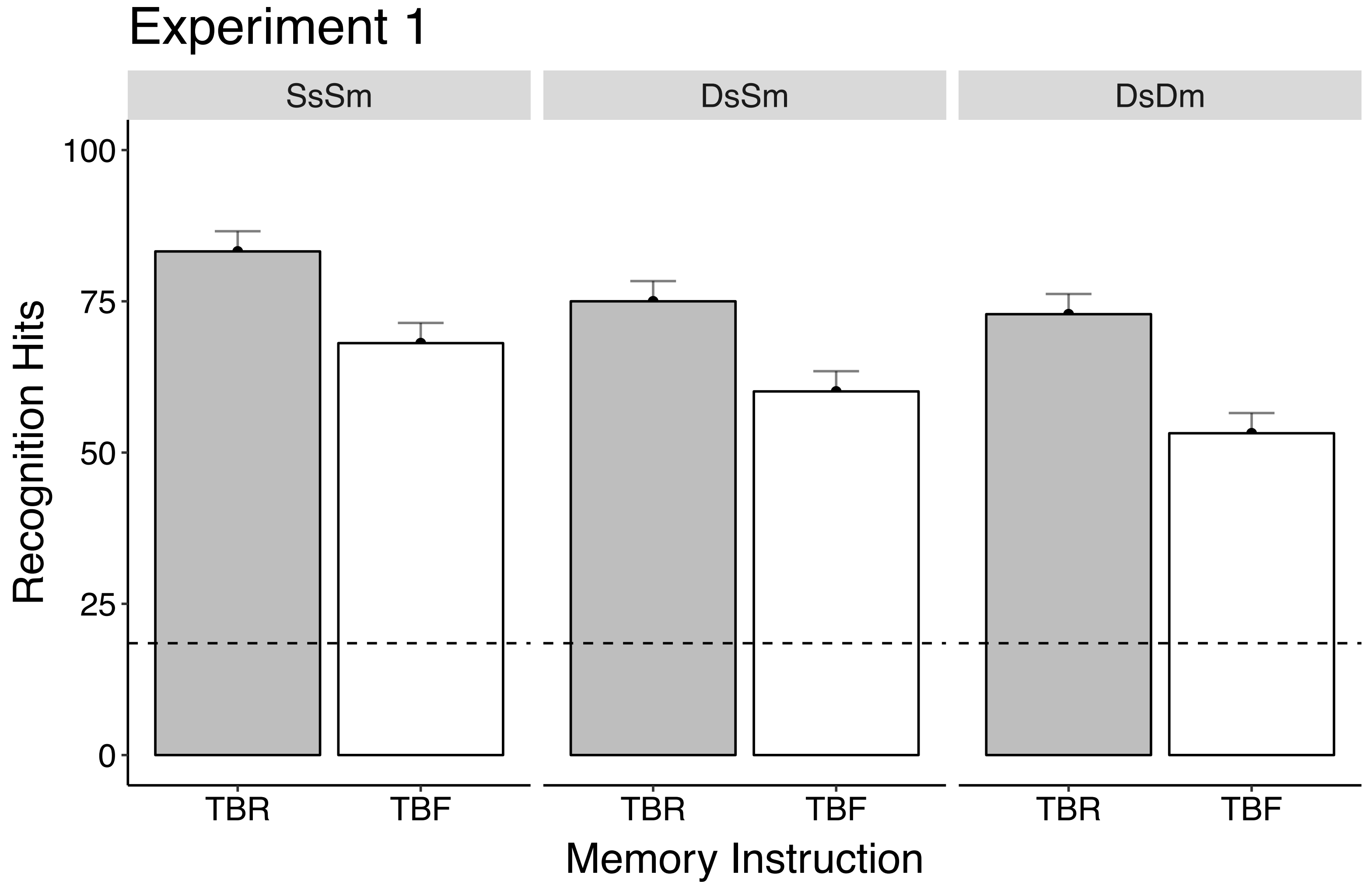
Figure 4

Recognition hits (%) on the yes-no recognition test in Experiment 3. As described in the text, conditions were initially dummy-coded according to the SsSm, DsDm, DsDm condition labels used to distinguish the different context conditions in Experiments 1 and 2. These labels were then changed to reflect the fact that the mock x-string "sentences" could not establish same or different meanings for the embedded words; instead, all that was manipulated was whether the embedding x-strings were the same (SsSm → Ss) or different (DsDm → Ds1, DsDm → Ds2) at test. The dashed horizontal line represents the overall false alarm rate to foils. Error bars depict Fisher's Least Significant Difference for the plotted effect

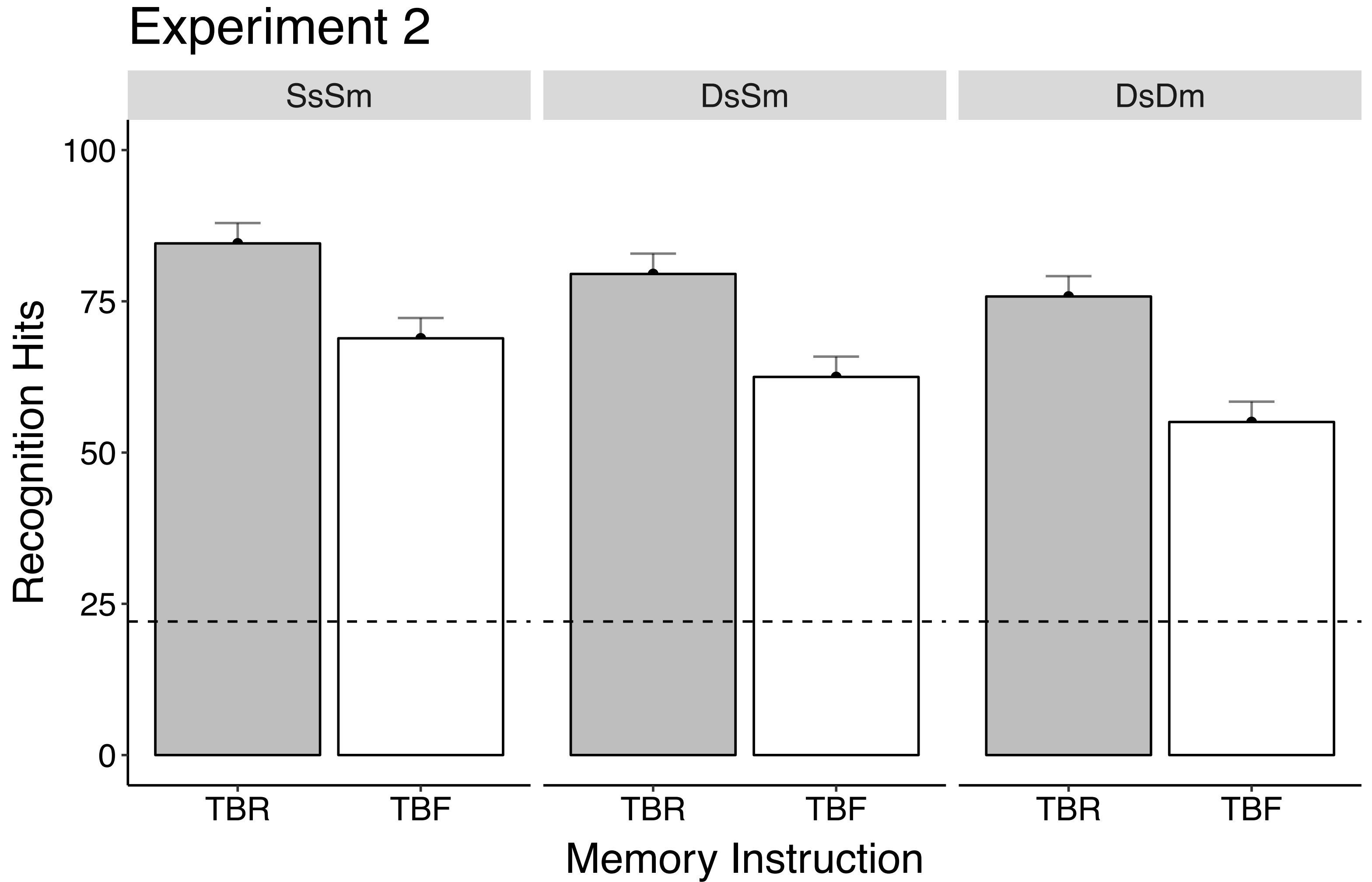
Figure 5

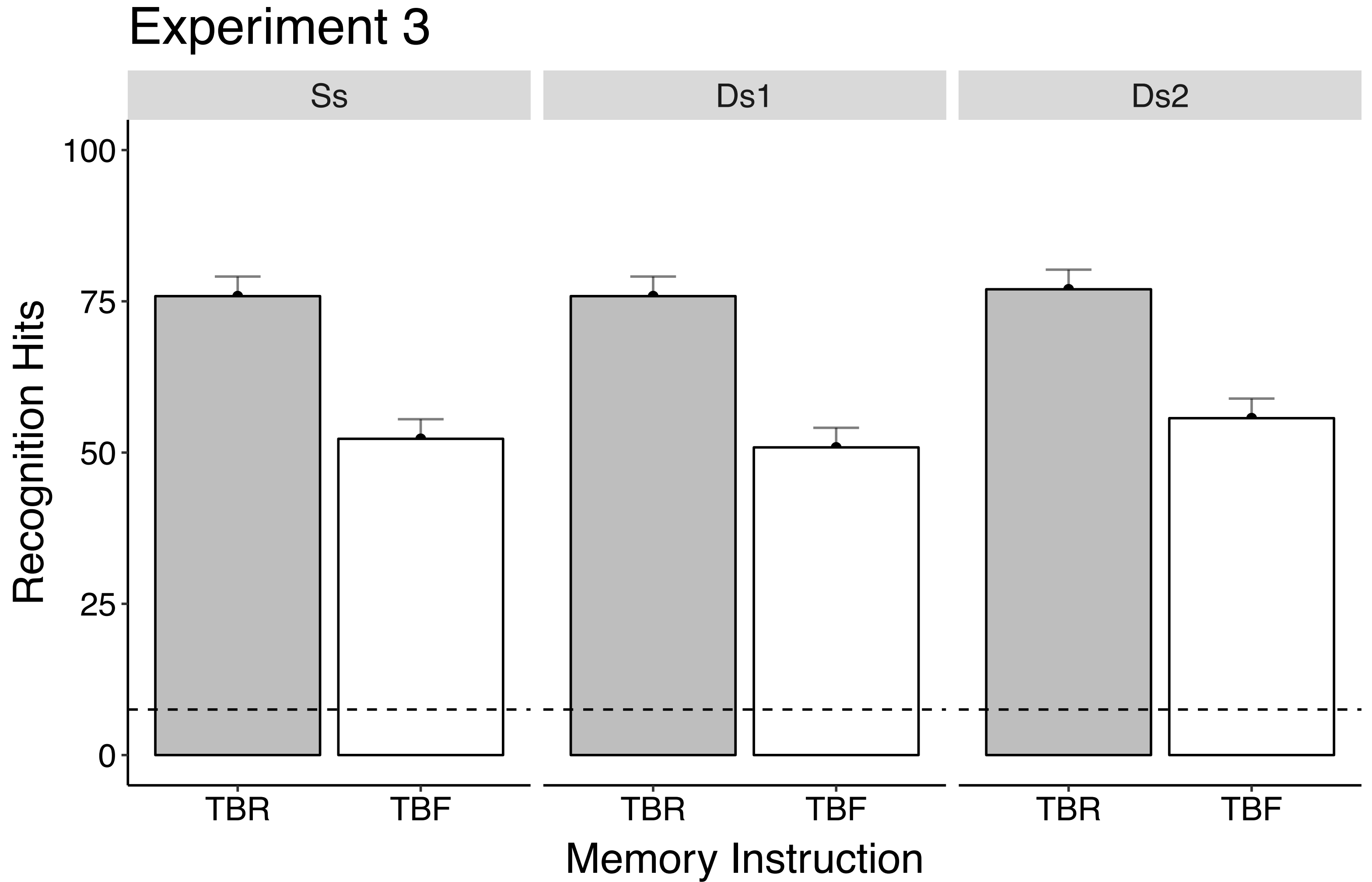
Recognition hits (%) on the yes-no recognition test in Experiment 4, as a function of memory instruction (TBR, TBF) and study-test context (SsSm=Same meaning, same sentence; DsSm = Same meaning, different sentence; DsDm = Different meaning, different sentence). The dashed horizontal line represents the overall false alarm rate to foils. Error bars depict Fisher's Least Significant Difference for the plotted effect.


Study
1,000 ms
+
3,000 ms
Drinks and snacks were available in the third coach of the train.
3,000 ms
Drinks and snacks were available in the third coach of the train.
500 ms
+
500 ms
RRRRRRRRRR
Recognition
Same Sentence
Same Meaning
Drinks and snacks were available in the third coach of the train.
Different Sentence
Same Meaning
The horse-drawn coach rattled noisily over the cobblestone drive.
Different Meaning
Peter sometimes felt like he could use a life coach.
or
The football coach was tough but the players liked her.

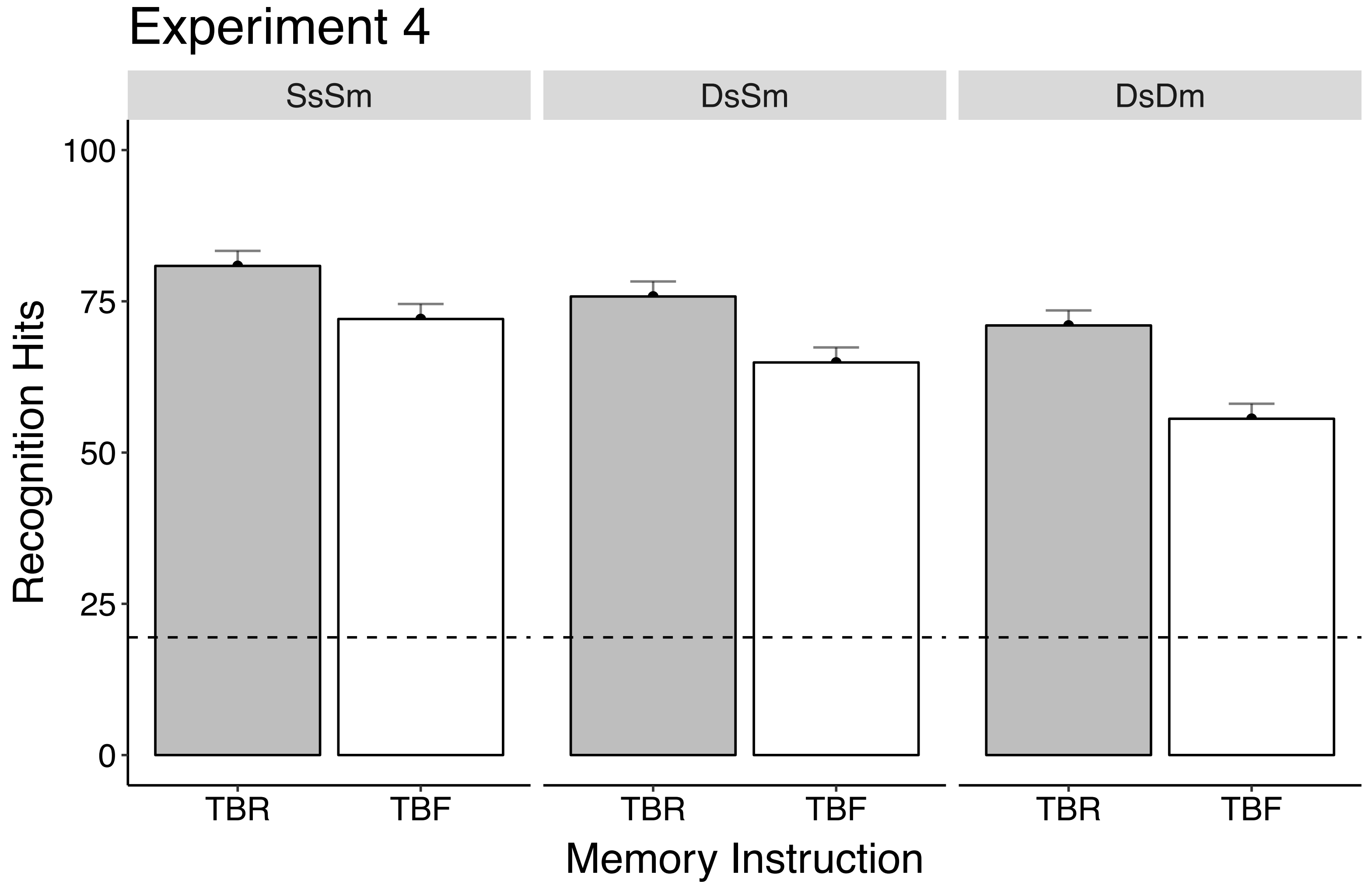


Experiment 1
SsSm
DsSm
DsDm
100
75
50
25
0
Recognition Hits
TBR
TBF
TBR
TBF
TBR
TBF
Memory Instruction


Experiment 2
SsSm
DsSm
DsDm
100
75
50
25
0
Recognition Hits
TBR
TBF
TBR
TBF
TBR
TBF
Memory Instruction



Experiment 3
Ss
Ds1
Ds2
100
75
50
25
0
Recognition Hits
TBR
TBF
TBR
TBF
TBR
TBF
Memory Instruction



Experiment 4
SsSm
DsSm
DsDm
Recognition Hits
100
75
50
25
0
TBR
TBF
TBR
TBF
TBR
TBF
Memory Instruction