Examining the lag effect under incidental encoding: Contributions of semantic priming and reminding

Dung C. Bui\textsuperscript{a}, Geoffrey B. Maddox\textsuperscript{a}, Fan Zou\textsuperscript{a} & Sandra S. Hale\textsuperscript{a}
\textsuperscript{a} Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA
Published online: 22 Apr 2014.
Examining the lag effect under incidental encoding: Contributions of semantic priming and reminding

Dung C. Bui, Geoffrey B. Maddox, Fan Zou, and Sandra S. Hale

Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA

Memory is better when repeated learning events are spaced than when they are massed (spacing effect), as well as when material is processed semantically than when it is processed graphemically (levels-of-processing effect). Examination of the relationship between levels of processing and spacing for both deeply and shallowly encoded items has shown a spacing effect for items processed deeply, but not shallowly. A semantic priming account of spacing was proposed to explain the interaction between levels of processing and spacing on memory. The current study manipulated levels of processing and the amount of spacing (lag) that occurred between repetitions of items that were incidentally encoded. Results from Experiments 1A and 1B revealed lag effects in test performance when items were deeply and shallowly encoded. Although these findings are inconsistent with a semantic priming account, they can be interpreted within a reminding account, which is explored in Experiment 2. Results from the second experiment indicate that bringing reminding under conscious control benefited items that were presented at a long lag but not at a shorter lag. Together, this study provides evidence that is difficult to accommodate with a semantic priming account of spacing and instead provides additional support for a reminding account suggesting that automatic and controlled processes may both underlie the reminding process.

Keywords: Spacing effect; Lag effect; Levels of processing; Semantic priming; Reminding.

One of the most robust findings in memory research is the long-term memory benefit obtained by separating repeated study events with intervening time or material relative to massing repeated study events (i.e., the spacing effect). Indeed, the spacing effect was first reported by Ebbinghaus (1885) and has since been observed across species, populations, and a range of experimental manipulations (see Crowder, 1976, Chapter 10; Delaney, Verkoeijen, & Spirgel, 2010; Dempster, 1996, for reviews). One recent meta-analysis (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006) revealed a nonmonotonic relationship between long-term memory performance and the spacing interval between repeated study events, a finding that is known as the lag effect (see also Glenberg, 1976).

Given that spacing and lag effects are both ubiquitous findings in the memory literature, rare failures to obtain these effects can provide unique leverage to uncover the mechanism(s) underlying the benefit of spaced study. One such failure—namely, the lack of a spacing effect when material is incidentally encoded at a shallow level—is the focus of the current study. A benefit of spacing is obtained less frequently with incidental encoding than with intentional encoding, and in instances where a spacing...
effect is obtained with incidental encoding, the magnitude of the effect is often smaller than the effect obtained under intentional encoding (e.g., Russo, Parkin, Taylor & Wilks, 1998; Shaughnessy, 1976). The current study manipulated levels of processing in an incidental encoding task during which items were repeated after a short or long lag. Before turning to the current set of experiments, however, we provide a brief review of past spacing effect studies that have used intentional encoding as a way of introducing some of the mechanisms proposed to account for the spacing effect. Then we consider past studies that have examined the spacing effect with the use of incidental encoding and discuss the implications of those results for mechanisms proposed to underlie the benefit of spacing.

Mechanisms proposed to underlie the spacing effect

Early accounts of the spacing effect focused on single mechanism explanations. For example, the deficient processing account (e.g., Jacoby, 1978; Rundus, 1971) suggested that massed presentations of material led to reduced processing on the item’s second presentation compared to when the second presentation was spaced. Furthermore, it has been suggested that deficient processing may occur automatically or can be consciously controlled. For instance, rehearsal of an item increases as the amount of spacing (i.e., lag) between repetitions increases (Rundus, 1971; see also Delaney & Verkoeijen, 2009). Similarly, study time for once-presented items, massed repetitions, and spaced repetitions increased when participants were allowed to control their study time (Zimmerman, 1975; see also Shaughnessy, Zimmerman, & Underwood, 1972).

With respect to automatic accounts of deficient processing, Jacoby (1978) had participants complete fragmented target words missing one letter (easy completion) or two letters (difficult completion). A surprise test revealed monotonically increasing performance as lag increased for fragments that were difficult to complete, whereas performance initially increased and then decreased as lag increased for easy fragments. Across both difficulty conditions, the important finding was that performance initially improved with increased lag, which Jacoby argued was due to the distinction of solving the fragment versus remembering the answer (i.e., long lag vs. short lag, respectively).

Subsequent accounts of the spacing effect have incorporated multiple mechanisms in order to explain all results. One account proposed by Greene (1989; see also Raaijmakers, 2003) suggested that study-phase retrieval (Thios & D’Agostino, 1976) and encoding variability (Estes, 1955a, 1955b) may optimally combine such that the second presentation of an item triggers retrieval of the item’s first presentation. When such retrieval is successful, long-term retention for this item is increased with longer spacing intervals due to increased encoding variability. In other words, repetitions would be encoded more variably when separated by greater amounts of time or intervening material, which in turn would lead to increased retrieval cues to be used on the final test.

More recently, the study-phase retrieval account has been revived in terms of a “reminding account” that incorporates a study-phase retrieval mechanism. However, it deviates from the accounts proposed by Greene (1989) and Raaijmakers (2003) in that it does not directly rely on encoding variability. This reminding account (Benjamin & Tullis, 2010) suggests that the second presentation of a repeated item has some capacity to remind participants of an earlier presentation of the item, which in turn benefits later memory for that item. Benjamin and Tullis (2010; see also Hintzman, 2004, 2008, 2010, 2011) have suggested that reminding occurs automatically, whereas results reported by Wahlheim, Maddox, and Jacoby (2014) suggest that reminding may occur automatically, but can also be brought under conscious control of the participant by instructing them to identify repetitions. In either case, it is suggested that successful reminding leads to an increase in performance over conditions in which items are repeated but reminding fails.

The spacing effect and intentionality of learning

As noted in the previous section, early mechanisms were singular in nature, but experimental
manipulations across numerous studies highlighted boundary conditions in which the spacing effect was not obtained. Indeed, the manipulation of intentionality in past studies has been useful in discriminating the degree to which various mechanisms can account for the spacing effect. Traditionally, incidental encoding studies have utilized an orienting task paired with a cover story to ensure that participants process the to-be-tested material without invoking any intentional learning processes. One common orienting task is to ask participants to rate stimuli on various dimensions. In one study, Craik and Tulving (1975) asked participants to process some dimension of the stimulus’s graphemic characteristics (“Is this word written in capital letters?”), phonemic characteristics (“Does this word rhyme with ___?”), or semantic meaning (“Would the word fit in the sentence___?”). These different questions were intended to manipulate the level of processing for each word such that structural, phonemic, and semantic questions corresponded to shallow, medium, and deep levels of encoding, respectively. Critically, semantic processing led to higher recall than phonemic condition, which in turn led to higher recall than the graphemic encoding. The authors took this pattern of results as support for the idea that the strength of the memory trace varies as a function of the depth of encoding.

If one considers studies that have compared intentional and incidental encoding in a spaced study paradigm, a spacing effect has often been observed in both encoding conditions. However, the effect is typically smaller for incidental learning than for intentional learning (e.g., Russo et al., 1998; Shaughnessy, 1976). Shaughnessy (1976) had participants rate words using two different scales or study the items intentionally before completing a free recall test. He found a larger spacing effect under intentional encoding than under incidental encoding. In a similar study, Russo et al. (1998) examined the spacing effect in recognition memory under intentional and incidental learning conditions. Although Russo et al. did not directly compare the size of the spacing effect under incidental and intentional learning, the effect was larger under intentional than under incidental learning. Another study reported by Verkoeijen, Rikers, and Schmidt (2005) repeated words at multiple lags, and recall results revealed a longer optimal lag (the lag at which performance is maximized) for intentional encoding than for incidental learning. However, the finding of a larger spacing effect in intentional encoding than in incidental encoding was still obtained at each condition’s optimal lag.

Taken together, evidence indicates that a spacing effect can be obtained under incidental encoding conditions. Importantly, such findings cannot be accommodated by a deficient processing mechanism that operates at a conscious level, because one would not expect differential rehearsal of items under incidental encoding. It is also the case that the spacing effect is observed even when each presentation of a given item is rated using different scales (e.g., Shaughnessy, 1976), which critically demonstrates that an automatic deficient processing account also cannot accommodate all situations in which a spacing effect is obtained. In this situation, using different scales would require that the item be rated on different dimensions and that a previous rating could not be retrieved and offered as a response on the item’s second presentation. Although controlled and automatic deficient processing accounts may still influence the magnitude of the spacing effect under intentional learning conditions, studies examining the spacing effect under incidental learning conditions nicely demonstrate how a failure to obtain the spacing effect can constrain possible underlying mechanisms. One limiting factor of previous studies, however, is that almost all have used incidental encoding conditions that emphasize deep levels of processing. Indeed, examination of the spacing effect under shallow levels of processing may also contribute to our understanding of the underlying mechanisms of the spacing effect. As such, the manipulation of levels of processing and lag under different incidental encoding conditions are considered next.

### The spacing effect and levels of processing

As noted above, several studies have obtained spacing effects with incidentally learned
information, but these studies all utilized deeper levels of processing for their rating tasks (Challis & Brodbeck, 1992; Greene, 1989; Rose & Rowe, 1976). In a study by Challis (1993), however, participants studied words under one of three learning conditions: in the intentional learning condition, they were simply told to study the words for a later test; in the incidental–graphemic (shallow) condition, participants made judgements on physical characteristics of each word; and in the incidental–semantic (deep) condition, participants made pleasantness ratings for each word. Words were presented either once or twice, with the repetition of words occurring after one of several possible lags. At a final test, participants were cued using words that were graphemically similar to the target word that was supposed to be recalled (e.g., FICKLE would be a cue for the target word FRECKLE). Results showed both spacing and lag effects in the intentional and incidental–semantic (deep) conditions, but there was no evidence of a significant spacing or lag effect in the incidental–graphemic condition.

Challis (1993) proposed a semantic priming account of spacing in which priming (the finding that presentation of a particular item facilitates responses on subsequent presentations of the same stimulus; Coane & Balota, 2010; Tulving, Schacter, & Stark, 1982) has a larger effect for massed items than for spaced items. If the effect of semantic priming decreases as the time between repetitions increases, one would expect less semantic processing of the repeated presentation when it is massed than when it is spaced. As a result, spaced presentations will receive more overall semantic processing than massed presentations. And because these spaced items are processed more than massed items, a spacing effect arises. With respect to incidental encoding with a shallow orienting task, one may not expect either a spacing or a lag effect, because no semantic processing or priming would occur (Russo et al., 1998).

However, using an incidental learning procedure, Russo et al. (1998) found spacing effects in a face recognition task under learning conditions that oriented participants to the structural (i.e., nonsemantic) features of the target items. This was followed up by Russo and Mammarella (2002), who also found a spacing effect for nonwords under incidental learning conditions that promoted nonsemantic (structural–perceptual) processing of the target items. Together, the findings of these two studies strongly suggest that the semantic priming account proposed by Challis (1993) cannot fully account for the spacing effect, as any kind of semantic processing of the target items was unlikely in the studies reported by Russo et al. (1998) and Russo and Mammarella (2002).

This led Russo et al. (1998) to incorporate the semantic priming account of spacing (Challis, 1993) in a transfer-appropriate processing approach (Morris, Bransford, & Franks, 1977) to memory, arguing for the congruency between the processes invoked on target items during study and the processes invoked on targets at the time of final test. In its simplest form, if the final test requires the type of processing that was necessary during the acquisition phase, then a lag effect should be observed as a result of deficient processing (i.e., semantic or structural priming). On one hand, if the type of processing that takes place during the test is primarily semantic in nature, then one would expect to observe a spacing effect when participants are also oriented towards the semantic features of the target items earlier during study (e.g., Challis, 1993). On the other hand, if the type of processing that takes place at the time of both study and test is nonsemantic in nature, then the spacing effect is thought to be due to some form of short-term perceptual priming mechanism. More specifically, Russo and colleagues (1998, 2002) argued that performance on a memory task for nonsemantically processed targets depends on the extent to which structural–perceptual information of a given target item is invoked. The structural–perceptual analysis that takes place during the first presentation of a target is thought to serve as a prime during the repeated presentation, which in turn facilitates the structural–perceptual processing of the second occurrence.

**Current study**

According to the accounts proposed by Challis (1993) and Russo and colleagues (1998, 2002),
spacing effects should not be observed in a semantic recall task when participants perform a nonsemantic orienting task (shallow levels of processing) during the learning phase. Indeed, the study by Challis (1993) demonstrates an interaction between lag and levels of processing, such that the spacing effect is not observed when items are processed at a shallow (nonsemantic) level during the acquisition phase. Although this finding lends support to the accounts put forth by Challis (1993) and Russo and colleagues (1998, 2002), the study by Challis (1993) provides the only evidence for this crucial prediction made with regards to the relationship between lag and levels of processing.

The goal of this study, therefore, is to conceptually replicate the findings by Challis (1993) and to assess the robustness of the interaction between levels of processing and the lag effect. To the extent that the findings of our study deviate from predictions made by Challis (1993) and Russo and colleagues (1998, 2002), our goal was to also explore other mechanisms of the spacing effect that may better accommodate our obtained results.

**EXPERIMENT 1A**

**Method**

**Participants and design**
Fifty undergraduate students from Washington University in St. Louis (25 females; \( M_{\text{age}} = 19.3 \) years, \( SD = 1.1 \)) participated in the study as partial course fulfilment. Lag (2 items vs. 12 items) was manipulated within participants, whereas level of processing (shallow vs. deep) was manipulated between participants.

**Materials**
The study list was composed of 96 trials. The first and last four trials of the study list were used as buffer items to account for any primacy and recency effects in recall. Of the remaining trials, 24 trials were filler items that were each presented once throughout the entire learning phase. The remaining trials were composed of 32 critical items, each of which was presented twice throughout the learning phase, half at lag 2 and the other half at lag 12. Filler items were included to ensure that average serial position of both presentations were equated between the lag conditions. All of the items were counterbalanced across the different lag and levels-of-processing conditions, and word length and frequency were equated across sets (\( ps > .30 \)).

Items in the shallow encoding condition were paired with one of two questions: “Does this item end in a vowel?” or “Does this item have an even number of letters?” Words in the deep encoding condition were paired with one of the two following questions: “Is this item animate?” or “Is this item larger than a shoebox?”. Each filler item was randomly assigned to one of the two questions from its corresponding levels-of-processing condition. The critical items were also paired with the questions such that they did not see the same question twice for each item. For example, if during the first presentation the participants were asked, “Does this item end with a vowel?” they would then be asked “Does this item have an even number of letters?” during the second presentation. For each word presented, participants were given 5 s to answer the levels-of-processing question, after which the next word was presented. Finally, after participants answered the levels-of-processing question, the word remained on the screen for the remainder of the 5 seconds.

The word stem completion test consisted of words all from the previous study list, presented one at a time in random order. The first two letters of each item were used as the stem and were followed by a line (e.g., TA____ would be the stem for the word TABLET), and the order of the stems was randomized. Additionally, each target word was uniquely cued by their first two letters. Participants were given a maximum of 20 s to respond for each trial.

**Procedure**
At the start of the experiment, participants were first asked to make judgements on a series of words that were presented individually on the computer (i.e., they were not informed that they would
be tested on the words). The participants were told that for each trial, they would see a word at the top of the screen with a yes/no question right below it. Their task was to type their response into a box that was provided. The incidental encoding phase was followed by a short retention interval, where participants completed simple maths problems for 30 s. Following the short retention interval, participants were given the word-stem completion test. The stems were presented one at a time on the computer, and participants were instructed to complete the stems with words they had previously rated in the judgement task.

**Results**

Performance on the word-stem completion task as a function of lag and retention interval is presented in Figure 1. A 2 (lag: long vs. short) × 2 (level of processing: shallow vs. deep) mixed-factor analysis of variance (ANOVA) revealed an effect of lag, $F(1, 48) = 32.85$, $p < .001$, $\eta_p^2 = .41$, such that items presented at a longer lag were better remembered on the final test than items presented at a short lag ($M_{\text{diff}} = .12$, 95% CI [.08, .17]). Additionally, there was an effect of levels of processing, $F(1, 48) = 5.42$, $p = .024$, $\eta_p^2 = .10$; not surprisingly, deep items were better recalled than shallow items ($M_{\text{diff}} = .09$, 95% CI [.01, .17]). Critically, no interaction was found between the two factors, $F(1, 48) = 1.71$, $p = .198$, $\eta_p^2 = .03$.

The results of this experiment replicate past findings of the spacing effect, as well as the levels-of-processing effect. More importantly, however, we were able to obtain a spacing effect for items under incidental–shallow encoding conditions. Given past failures to find a lag effect under shallow–incidental processing, the current experiment was replicated with a more generalizable sample.

**EXPERIMENT 1B**

The goal of Experiment 1B was to replicate the findings of Experiment 1A by using the same method and procedure but with an internet-based sample.
Method

Thirty participants (20 females; $M_{\text{age}} = 33.6$ years, $SD = 10.7$), were recruited from Amazon’s Mechanical Turk (see Mason & Suri, 2012, for a recent overview) web site to take part in this study for monetary compensation ($0.30). Lag (2 items vs. 12 items) was manipulated within participants, whereas level of processing (shallow vs. deep) was manipulated between participants. The materials and procedure used for this experiment were identical to those used in Experiment 1A.

Results and discussion

Performance on the final test as a function of lag and retention interval is presented in Figure 2. A 2 (levels of processing) × 2 (lag) mixed-factor ANOVA revealed an effect of lag, $F(1, 28) = 11.71, p = .002, \eta_p^2 = .30$, such that long lag items were better recalled than short lag items ($M_{\text{diff}} = .10, 95\% \text{ CI} [.04, .16]$). Additionally, an effect of levels-of-processing condition was found, $F(1, 28) = 4.54, p = .042, \eta_p^2 = .14$; not surprisingly, deeply processed items were better recalled than shallowly processed items ($M_{\text{diff}} = .12, 95\% \text{ CI} [.01, .25]$). Finally, there was no interaction between the two factors, $F(1, 28) = 0.77, p = .39, \eta_p^2 = .03$.

Our attempt to replicate Experiment 1A with a more generalizable sample was successful, as results indicated that the spacing effect can be obtained under incidental–deep encoding conditions as well as under incidental–shallow encoding conditions.

Taken together, the results of the present study suggest that the magnitude of the lag effect is similar for shallow and deep levels of processing. This finding argues against the accounts of Challis (1993) and of Russo and colleagues (1998, 2002), who would predict an interaction, such that spacing effects are observed for items processed deeply (semantically), but not shallowly (nonsensematic). It is unclear as to why our results differed from that of Challis (1993), though it is worth pointing out that the lags used in Experiment 1A and 1B (2, 12) were different from those used by Challis (0, 10, 30). Although our experimental design...
precludes us from making direct comparisons, and similarly precludes us from making any definitive statements, it is important to note that neither Challis (1993) nor Russo et al. (1998) would predict our observed spacing effect under shallow levels of processing, even with our selected lags. Indeed, our results in Experiments 1A and 1B seem to indicate that the semantic priming account (Challis, 1993) and semantic priming/transfer-appropriate processing account (Russo et al., 1998) cannot fully accommodate the spacing effect.

The results reported thus far are, however, consistent with an automatic reminding account of spacing and lag effects (e.g., Benjamin & Tullis, 2010), an account that we touched on briefly earlier during the introduction of this paper. Therefore, the goal of this next experiment (Experiment 2) was to explore the extent to which the reminding account can accommodate the results we have obtained in the previous two experiments (1A and 1B), and how different processes (automatic vs. controlled) may influence successful reminding in our paradigm.

**EXPERIMENT 2**

According to the reminding account of spacing and lag effects (Benjamin & Tullis, 2010), the second presentation of a repeated item is thought to remind participants of an earlier presentation of the item, which then benefits later memory for that item. In the context of a spacing paradigm, reminding is less likely to be successful for items that are separated by longer lags than by shorter lags. However, a number of studies suggest that retrieval serves as a learning event (e.g., Carrier & Pashler, 1992; Karpicke & Roediger, 2006) and that the more difficult it is to retrieve a particular piece of information, the more beneficial the retrieval event is for strengthening memory (e.g., Bjork, 1994). Thus, in terms of a spacing paradigm, retrieval of the first presentation of an item becomes increasingly difficult as the lag increases between presentations, but when successful, retrieval results in better memory for spaced items than for massed items. More specifically, memory increases as the lag between repetitions of an item increases. As a result, the extent or quality to which items are initially encoded should modulate the optimal lag at which retrieval is successful, and difficulty is maximized.

With regards to the results obtained in Experiments 1A and 1B, an automatic reminding account of spacing would predict that, provided that the lags are conducive for successful reminding, lag effects should be obtained even when material is incidentally encoded with shallow processing. Indeed, our results from those two studies yielded a spacing effect for shallowly processed items. Of course, it may not always be the case that reminding is successful for participants in this task. Thus, it is important to consider recent evidence that supports the distinction between automatic and controlled reminding and how that may influence results in the current paradigm.

Although the reminding account of spacing has been largely conceptualized as a spontaneous process (Benjamin & Tullis, 2010), there is evidence by Wahlheim et al. (2014) to suggest that although reminding may be automatic in nature, it can also be brought under conscious control of the participant. In instances of both automatic and controlled reminding, successful reminding should result in improved memory performance over conditions in which items are repeated but reminding fails. Furthermore, it is possible that controlled reminding may be used to help participants retrieve repetition events in instances where automatic reminding fails, and participants are unable to spontaneously make contact between repetitions. Indeed, results by Wahlheim et al. (2014) show that when reminding is brought under conscious control, enhanced spacing effects can be obtained, such that reminding can occur following longer lags. In such cases, the relationship between lag and reminding highlights the boundary conditions that should be considered to ensure successful reminding. In instances where reminding is unsuccessful, spacing effects are not expected to be observed.

To the extent that both automatic and controlled reminding processes play a role in modulating the
spatial effect (e.g., Wahlheim et al., 2014), it is important to understand how bringing reminding under conscious control by participants in the present study influences the magnitude of the spacing effect. Indeed, when participants are unable to make spontaneous contact between repetitions in the current study, it may be the case that bringing reminding under conscious control may help participants make contact between two repeated presentations. In turn, this should allow us to observe stronger spacing effects in the current incidental encoding paradigm. Therefore, the goal of this experiment was to explore the effects of a controlled reminding procedure in an incidental learning paradigm by including a condition in which participants made a reminding judgement on each trial of the incidental learning task.

Results and discussion

Performance on the word-stem completion task as a function of lag, levels of processing, and reminding is presented in Figure 3. A 2 (lag: long vs. short) × 2 (level of processing: shallow vs. deep) × 2 (reminding: automatic vs. controlled) mixed-design ANOVA revealed a main effect of lag, $F(1, 64) = 41.98, p < .001$, $\eta_p^2 = .40$, such that items presented after a long lag were better recalled than items presented after a short lag ($M_{diff} = .11, 95\% CI [.07, .14]$). We also obtained an effect of levels of processing, $F(1, 64) = 12.03, p < .001$, $\eta_p^2 = .16$, indicating that items were better recalled if they were processed deeply than if they were processed shallowly ($M_{diff} = .10, 95\% CI [.04, .15]$). Finally, there was an effect of reminding, $F(1, 64) = 5.87, p = .018$, $\eta_p^2 = .08$, indicating that recall was better in the controlled condition than in the automatic condition ($M_{diff} = .07, 95\% CI [.01, .12]$).

Consistent with the previous experiments, no interaction was found between lag and levels of processing, $F(1, 64) = .96, p = .332$, $\eta_p^2 = .02$. However, the interaction between lag and reminding was significant, $F(1, 64) = 5.71, p = .020$, $\eta_p^2 = .08$. To further explore the lag by reminding interaction, follow-up tests showed that performance was better on long lag items for the controlled condition ($M = .29$) than for the automatic condition ($M = .18$), $t(66) = 2.98, p = .004, d = .72, 95\% CI_{mean difference} [.04, .18]$. However, performance on short lag items did not differ between the controlled condition ($M = .14$) and the automatic condition ($M = .11$), $t(66) = .91, p = .369, d = .22, 95\% CI_{mean difference} [−.04, .09]$.

Of course, it is also important to consider how often repetitions were successfully detected. This was done by examining responses on the second presentation of an item and whether participants correctly responded yes to whether they had previously seen the presented word on the screen. Responding yes in these instances indicated that participants were able to detect an item’s repetition and thus make contact between the two presentations for successful reminding. Examination of these reminding judgements revealed a high

Method

Participants and design

A total of 68 participants (39 females, 27 males; $M_{age} = 30.1$ years, $SD = 10.1$), were recruited from Amazon’s Mechanical Turk web site to take part in this study for monetary compensation ($0.30). Lag (2 items vs. 12 items) was manipulated within participants, whereas level of processing (shallow vs. deep) and reminding judgement (automatic vs. controlled) were manipulated between participants.

Materials and procedure

For participants in the automatic reminding condition, the materials and procedure used for this experiment were similar to those used in Experiment 1A and Experiment 1B. Two changes were implemented for participants in the controlled reminding condition. After participants answered the levels-of-processing question, they were then asked “Did you previously make a judgement on this word?” Participants responded to this reminding question with a yes/no response. Since participants had to respond to two different questions for each word, presentation time for each word was increased to 6 s.
proportion of correct responses ($M_{\text{shallow-short lag}} = .90, SD = .30$; $M_{\text{shallow-long lag}} = .88, SD = .32$; $M_{\text{deep-short lag}} = .89, SD = .34$; $M_{\text{deep-long lag}} = .88, SD = .32$), indicating that this looking-back procedure was able to bring reminding under some form of conscious control. This high rate of successful reminding, particularly for items presented after a long lag, is one possible reason as to why performance for items presented after a long lag was selectively boosted when a controlled reminding procedure was introduced.

We were once again able to obtain a lag effect as well as a levels-of-processing effect in Experiment 2. As with Experiments 1A and 1B, we were able to demonstrate that lag effects with shallowly processed items can be obtained. Furthermore, the interaction between lag and reminding suggests that making reminding judgements had a selective benefit for items presented after a long lag and not for items presented after a short lag. This makes sense given that spontaneous reminding should be easier and relatively more successful after short lags than after long lags. Since the items in the long lag condition benefited from a consciously controlled reminding procedure, this seems to suggest that automatic reminding may not always be successful in these instances. When such reminding fails, a consciously controlled reminding procedure can help participants make contact between repetitions, thus allowing them to benefit more from spacing at longer lags.

One might be concerned that adding a controlled reminding judgement during the rating task may have induced an intentional learning strategy. Although there may have been a qualitative shift in processing during acquisition to allow for detection of repetitions, if an intentional learning strategy had been induced one would predict significant increases in performance across both lag conditions relative to performance in the incidental (without reminding judgement) condition. However, the benefit of controlled reminding was isolated to the long lag condition, and closer inspection of the short lag condition did not show a significant effect of reminding condition on performance for either shallow ($t = 0.65$, $p = .520$, $d = 0.23$, 95% CI mean difference $[-.05, .09]$) or deep ($t = 0.70$, $p = .490$, $d = 0.25$, 95% CI mean difference $[-.07, .15]$) levels of processing.

It is worth pointing out that the lack of differences between reminding conditions for short lag items in the shallow condition may be due to issues of scaling in performance. That is, performance may be so low for short-lag, shallow items in the automatic

![Figure 3. Proportion of correct responses on final test as a function of lag, levels of processing, and reminding for Experiment 2. Error bars represent standard error of the mean.](image-url)
reminding condition \((M = .07)\) that floor effects may mask any differences between the two reminding conditions. Nonetheless, if the controlled condition did induce any intentional learning strategies, we would have also expected to see differences between reminding conditions in the deep processing condition where there was less of a concern for any floor effects (since performance was higher than that for shallowly processed items). Past research suggests that intentional learning may not increase performance above and beyond incidental deep processing (e.g., Challis, 1993), but switching from incidental shallow processing to intentional processing should benefit final test performance. The lack of a difference between conditions for short lag items (particularly in the deep processing condition) in the present experiment provide evidence that adding the reminding judgement in the current experiment did not induce intentional learning.

One final possible concern worth noting in this experiment is that participants in the controlled condition answered two questions per trial (the processing question and whether they remembered seeing the item from a previous trial), as opposed to only one question in the automatic condition. If answering more questions about a target item enhances memory for it at a later test, it is possible that the improvement in performance observed for long lag items in the controlled condition compared to the automatic condition is due to the additional processing received with two questions, and not necessarily to reminding. However, if this improvement was due primarily to this additional processing, the benefit observed for items in the long lag condition should be observed for short lag items as well. This was not the case as indicated by the interaction reported in the Results section. More specifically, memory for short lag items in the controlled condition was no different from that for the items in the automatic condition, despite participants in the former being presented with more questions. Conversely, there was an effect of reminding judgement on performance for long lag items. Thus, an “additional processing” explanation has challenges accounting for the current results, and although our results cannot speak to the possible contributions of any other additional processing, it does suggest that reminding is playing a role in the controlled condition.

**Meta-analysis**

Across all three experiments, we reliably demonstrated robust effects of lag and levels of processing, and, more importantly, we showed that those two factors do not interact with one another to influence recall. To further demonstrate the robustness of these effects, we submitted the three experiments reported in this study to a random-effects meta-analysis, which places more emphasis on estimation techniques. This approach, advocated by Cumming (2012, 2013), is in contrast to null-hypothesis significance testing, which may sometimes distort the results of a meta-analysis (Cumming, 2013). Instead, a meta-analysis approach provides a more precise estimation of the observed effects than any of the three experiments on its own and allows us to more carefully examine how robust our findings across the three experiments are.

To do such a meta-analysis, we included all participants from Experiment 1A \((n = 50)\) and 1B \((n = 30)\), as well as participants in the automatic condition in Experiment 2 \((n = 34)\). This method yielded a medium effect size of lag \((\text{Cohen’s weighted average } d = 0.64, \ SD = 0.01, 95\% \ CI [0.48, 0.79])\), a large effect of levels of processing \((\text{Cohen’s weighted average } d = 0.81, \ SD = 0.02, 95\% \ CI [0.57, 1.04])\), and no significant effect size for the interaction \((\text{Cohen’s weighted average } d = 0.01, \ SD = 0.03, 95\% \ CI [-0.32, 0.33])\). Tests for homogeneity of effect sizes indicated that the effect sizes across the three experiments did not significantly differ with regards to lag \((Q_2 = 0.06, \ p = .971)\), levels of processing \((Q_2 = 0.86, \ p = .650)\), and the interaction \((Q_2 = 1.62, \ p = .450)\). Taken together, our study demonstrates not only effects of lag and levels of processing (as well as no interaction between the two factors), but also that these effect sizes are stable across the three experiments.

**GENERAL DISCUSSION**

The current study examined the spacing effect under incidental learning conditions by manipulating
levels of processing. Using two different samples, Experiments 1A and 1B examined the semantic priming account (Challis, 1993) and transfer-appropriate processing account (Russo et al., 1998; Russo & Mammarella, 2002) of the spacing effect, and results from each study revealed spacing effects for items encoded under deep levels of processing as well as shallow levels of processing. The role of a consciously controlled reminding procedure was explored in Experiment 2, and results revealed that items presented after a long lag benefited from consciously controlled reminding, whereas items presented after a short lag did not.

Our results extend and further qualify past research that showed spacing effects under incidental–deep but not under incidental–shallow encoding conditions (Challis, 1993). Although these previous results generally provide support for a semantic priming account of the spacing effect as Challis (1993) argued, they present issues for a reminding account of spacing that relies on automatic processes (e.g., Benjamin & Tullis, 2010). In particular, a reminding account in which automatic processes underlie the spacing benefit predicts that spontaneous contact between repetitions should occur regardless of intentionality and level of processing if the repetitions are presented at lags that allow for successful reminding. Indeed, the results from both Experiments 1A and 1B prove to be difficult for a semantic priming account to accommodate but are consistent with the reminding account viewpoint. In particular, spacing effects were obtained under incidental learning conditions for both shallowly and deeply encoded items, a pattern that cannot be accounted for by a semantic priming account because orienting toward features of a stimulus that are not semantic in nature should not lend themselves to semantic priming. Additionally, a spacing effect with shallow levels of processing cannot be accounted for by the semantic priming/transfer-appropriate processing account (Russo et al., 1998) if one assumes that the processes invoked at the time of test in our study were semantic in nature. In this instance, Russo et al. (1998) would predict that the incongruency of processes during the test (semantic) and of processes invoked during the learning phase (nonsemantic) should not yield a spacing effect.

The results of Experiment 2 demonstrated that a controlled reminding procedure (Wahlheim et al., 2014) can be used to amplify the spacing effect. This is consistent with the idea that when spontaneous reminding fails, and participants are unable to make contact between repetitions, bringing reminding under conscious retrieval may help participants retrieve repetition events. With respect to Experiment 2, it is fair to assume that spontaneous contact between repetitions is more difficult and thus more likely to fail when items are separated by a long lag than when they are separated by a short lag. In such cases, it is not surprising to see that a controlled reminding procedure benefits repetitions separated by a long lag more than items separated by a short lag (conditions in which reminding is less likely and more likely to be successful, respectively). However, regardless of whether the reminding is taking place on an automatic or controlled level, successful reminding should result in improved memory performance over conditions in which reminding fails.

If the benefit of spacing is dependent on successful reminding, the likelihood of reminding should depend on the quality of encoding on the initial presentation of an item. Such a distinction highlights the role of cognitive abilities in the spacing effect and how individual differences may mediate the magnitude of an observed spacing effect. Work by Bui, Maddox, and Balota (2013) examined the roles of task difficulty in between repetitions and individual differences in working memory in a spaced repetition paradigm, both of which were believed to modulate successful reminding.

Individuals with high working memory ability benefited more from repeated study than those with lower working memory ability, regardless of task difficulty. More importantly, a crossover interaction was observed between working memory ability and task difficulty, such that low-working-memory individuals benefited more when the task difficulty was easy than difficult, but individuals with high-working-memory ability produced just the opposite effect. The results suggest that the
likelihood of successful reminding in certain contexts may vary among individuals, and they highlight the importance of individual differences in cognitive ability in optimizing the benefits of spaced learning.

It is also important to understand the results of the current study with respect to the role of retention interval. Indeed, studies have demonstrated that the lag effect tends to manifest over longer retention intervals. Specifically, Peterson, Wampler, Kirkpatrick, and Saltzman (1963) reported a significant lag by retention interval interaction in which massed retrieval produced a benefit in performance over spaced retrieval on an immediate test, but following a longer retention interval, the typical benefit in spacing was observed over massed retrieval (see also Balota, Duchek, & Paullin, 1989; Rawson & Kintsch, 2005). This lag by retention interval interaction was also reported by Cepeda et al. (2006) in their meta-analysis, and the critical finding is that the magnitude of the spacing effect increases as retention interval increases.

This interaction between lag and retention interval is of potential importance with regards to the current study, as it can provide insight as to whether the spacing effect observed with both shallowly processed items and deeply processed items are qualitatively different from one another. Indeed, any qualitative differences may be manifested over longer retention intervals, where one may see stronger spacing effects for items processed in one way over the other. Of course, one methodological challenge in extending a retention interval beyond the 30-s interval used in the current study is to ensure that performance is above floor. If future studies examining the relationship between lag and levels of processing as a function of retention interval can be mindful of this potential problem, these studies can help further our understanding of the mechanisms underlying the spacing and lag effect.

In short, the current study obtained a spacing effect in incidental learning when processing was manipulated across deep and shallow levels. This finding diverges from past research (e.g., Challis, 1993) and critically provides additional support for the reminding account of the spacing effect. Indeed, the benefit of spacing appears to be influenced by the extent to which the second presentation of an item reminds the participant of that item’s first occurrence. Moreover, in contrast with situations in which reminding occurs automatically, our results are consistent with recent research that demonstrates that such reminding can be brought under control of the participant, and such consciously controlled reminding can benefit performance in conditions when lag may otherwise be too long to facilitate high levels of successful, automatic reminding.

Original manuscript received 4 September 2013
Accepted revision received 16 December 2013
First published online 23 April 2014

REFERENCES


