SPACING AND LAG EFFECTS IN RECOGNITION MEMORY:

TIME VERSUS INTERVENING ITEMS

by

Matthew David Shively

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Psychology

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2007

APPROVAL

of a thesis submitted by

Matthew David Shively

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Richard Block, Ph.D.

Approved for the Department of Psychology

Richard Block, Ph.D.

Approved for the Division of Graduate Education

Carl A. Fox, Ph.D.

STATEMENT OF PERMISSION TO USE

Matthew David Shively

April 2007

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Spacing and lag effects both refer to abundant findings that memory is enhanced when repeated items are spaced. Several researchers have realized the difficulty of explaining these effects using only one theory, and, therefore posited varying dual-process models. It is also unclear if there is a limit to the increase in memory performance due to increased lag. This study sought to understand how stimulus type influences spacing and lag effects, limits of the lag effect, and the importance of time and items in creating these effects. Experiment 1 found a unique spacing effect and lag effect. Experiment 2 found no spacing effect, yet a lag effect was found. Both time and items are important in generating spacing and lag effects.

INTRODUCTION

The spacing effect is a well known and heavily researched phenomenon. More than 300 empirical studies have been published relating to the spacing effect (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Generally speaking, the spacing effect occurs when items that are repeated following intervening items are better remembered than items that are immediately repeated. The term *spacing effect* refers to an increase in memory performance as a result of having one or more items presented between repeated presentations of a target item (Kahana & Howard, 2005). A closely related effect of interest to research on the spacing effect is called the *lag effect*. The lag effect occurs when memory performance is a monotonically increasing (and usually also decelerating) function of the number of intervening items (i.e., the lag) between repeated items (Melton, 1970; Murdock, 1974).

The spacing effect is a robust effect that has been found using several different kinds of paradigms and stimuli. It has been observed using paired-associate learning, free recall, recognition memory, and distractor tasks (Hintzman, 1974). Stimuli used in spacing effect research have included nonsense syllables, words, sentences, and pictures (Hintzman, 1974; Janiszewski, Noel, & Sawyer, 2003). Adding to the complexity of using different paradigms and stimuli, researchers have conducted experiments under different encoding conditions by using incidental and intentional instructions (Greene, 1989, 1990; Paivio, 1974; Toppino & Bloom, 2002).

Many researchers have investigated the processes that contribute to spacing and lag effects. In his seminal work on the spacing effect, Hintzman (1974) discussed several types of processes that could produce spacing and lag effects. He separated these processes into either automatic or controlled processes. Hintzman's automatic processes include consolidation, habituation-recovery, and encoding variability. Controlled processes include voluntary attention and voluntary rehearsal.

Consolidation refers specifically to an encoding process by which stimuli are encoded into long-term memory from short-term memory. Once a trace exists in long-term memory, it becomes more easily accessible and therefore easier to retrieve a prior instance of a stimulus presentation. Hintzman (1974) noted that there is scant empirical evidence to support this theory. The habituation-recovery theory, like the consolidation theory, seeks to explain the spacing effect as a result of deficient processing of the second presentation of a stimulus. In this account, a presentation of a stimulus will activate a trace of the stimulus, which much like a neuron cannot be activated again for a certain period of time. If a second presentation is shown within the so-called *refractory period*, it will not receive the same amount of processing as if the presentation occurred after such a period. In contrast to the consolidation and habituation theories, the encoding variability theory predicts that the encoding of an item's second presentation is enhanced as spacing increases due to an inherent increase in the possible number of ways that the stimulus can be encoded. One example is variable encoding of word stimuli; as words are presented more than once, it is possible that more than one semantic meaning associated with the stimulus word is activated.

Aside from automatic process theories, Hintzman (1974) also suggested that voluntary processes could occur. Similar to the consolidation and habituation theories, the voluntary attention theory suggests that encoding of a second presentation of a stimulus is deficient. In the case of voluntary attention, this occurs because a participant does not allocate sufficient attention to the second presentation. This lack of attention could result from a participant's belief or expectation that the item is already memorable, a participant's lack of interest, or some similar voluntary process. Like the encoding variability theory, the voluntary rehearsal theory predicts enhanced encoding of a repetition as lag increases. Hintzman postulated that once a stimulus is first presented, it could be rehearsed and maintained in short-term memory only for a finite duration. If the first presentation of a repeated item were encoded into long-term memory, then a second presentation (occurring after the first presentation is encoded) would activate the trace encoded into long-term memory and also be rehearsed, thereby increasing the total rehearsal time for the repeated stimulus.

In a series of studies, Hintzman sought evidence to support or weaken theories of the spacing effect. For example, Hintzman, Summers, Eki, and Moore (1975) found some evidence against the voluntary attention theory. In their study, attention was manipulated by incentive and instructions. Neither manipulation of voluntary attention affected the spacing effect. Hintzman and Rogers (1973) also found weak evidence for the voluntary rehearsal theory. To test this theory, they used stimuli that theoretically cannot be rehearsed: pictures showing complex visual scenes (Shaffer & Shiffrin, 1972). They found the same pattern for the spacing effect as when stimuli included materials that could be rehearsed. They concluded that rehearsal has little effect on the spacing effect. To complicate matters, research also showed a spacing effect for repeated pictures using incidental conditions, but not for repeated words (Paivio, 1974). Finally, Hintzman, Summers, and Block (1975b) observed that predictions associated with the habituation-recovery hypothesis did not stand up to empirical test. This observation is in contrast to Hintzman's (1974) postulation that the habituation-recovery hypothesis is a strong candidate to explain the spacing effect, and further supported by Paivio (1974). Also noted by Hintzman et al. was an observation relevant to the present study—time between the offset of the first presentation of a stimulus and the onset of a second stimulus is an important factor in producing the spacing effect.

The observations by Hintzman and Rogers (1973) and Hintzman et al. (1975b) that time is an important factor in explaining the spacing effect seems to go against work by Glanzer (1969), who suggested that an increase in number of items (what he called distance) was the important manipulation which created the enhanced retrieval of the second presentation. Glanzer noted two possible reasons for this finding, which were roughly similar to Hintzman's (1974) proposed consolidation and encoding variability theories. Waugh and Norman (1968) also suggested that number of intervening items are more important than time in explaining interference in memory tasks. A recent meta-analysis supports the idea that time—specifically, interstimulus interval (ISI) and retention interval—is an important factor (Cepeda et al., 2006). Using a judgment of recency task, Hintzman (2004b) concluded that the amount of time between two items is what drives judgments of recency. The vast majority of research on the spacing effect in the last 30 years has focused more heavily on which processes contribute to the spacing and lag effects, while little research attention has been given to discovering which is more important to the these effects: time, items, or both?

Several researchers have suggested dual-process models to explain the spacing effect after years of not finding conclusive evidence for any one theory. One of the more interesting, albeit controversial, dual-process models comes from Greene (1989). Greene proposed that different processes affect memory performance differently for cued-memory tasks (e.g., recognition) and uncued-memory tasks (e.g., free recall). According to Greene, individuals engaging in cued memory tasks are more likely to rely on stimulus familiarity, at least insofar as participants are familiar with the experimenter supplied cues in such tasks. The initial familiarity is critical for a participant to determine the amount of rehearsal necessary to remember the stimulus. If a stimulus is presented twice within a short interval, a participant may be less likely to rehearse the stimulus because it seems familiar at that particular point. When repeated stimuli are presented further apart, familiarity decreases, thus increasing the likelihood that a participant will engage in additional rehearsal of the stimulus. This increased rehearsal also increases familiarity. Greene called this attenuation in attention due to familiarity a *deficient processing theory*. In contrast to Hintzman (1974), this deficient processing theory is a result of voluntary processing. This deficient processing theory is also similar in some ways to Hintzman's voluntary rehearsal theory in that when a repetition occurs, it is the trace of the repetition (e.g., second presentation) that is enhanced.

Although familiarity is an important factor for recognition tasks, contextual variability is important in free recall tasks. Greene argued that participants engaging in free recall rely heavily on associations between items, and such associations are strengthened by the increased number of contextual cues associated with longer lags. Temporal cues are one kind of contextual cues, and include intrinsic cues such as time between presentations. Greene called this theory a *study-phase retrieval* theory, using an earlier term in the literature (Hintzman, Summers, & Block, 1975a). More recently, Hintzman (2004a) called this process *recursive reminding*.

Aside from the differential effects of these processes on free recall and cued memory tasks, Greene further predicted that intentionality of encoding influences the spacing effect. Intentionality refers to the instructions given to participants before they study a list of to-be-remembered items. Intentional instructions are when participants are told to actively attend to and remember the items they see. Incidental instructions are instructions that do not reveal the memory aspect of the experiment. This may be done by administering a cover story or cover task. If, as Greene suggested, retrieval in cued-recall tasks depends on voluntary rehearsal, one would not expect to find a spacing effect in an incidental condition, in which participants are not inclined to rehearse. Moreover, if memory performance in a free recall task depends on an automatic process such as recursive reminding, one would expect to observe spacing and lag effects in both intentional and incidental conditions. These predicted patterns are exactly what Greene observed. Spacing effects were observed in intentional conditions across cued and free recall memory tasks. In the incidental-memory condition, a spacing effect was observed on the free recall task but on any cued-memory task, including a recognition-memory test. Greene (1990) replicated these findings in a second series of experiments.

Toppino and Bloom (2002) called into question Greene's (1989, 1990) findings. They replicated Greene's (1989) study, with the key exception that they controlled for serial position effects. Toppino and Bloom argued that Greene failed to control for such effects, and as a result Greene's findings should be viewed as specious and spurious. Toppino and Bloom found no spacing effect in free recall tasks under intentional conditions. Moreover, Toppino and Bloom found that memory was an inverted-U shaped function of lag. They showed that at an ISI of about 25 s the mean percentage of items correctly recalled on a free recall task peaked, and beyond that it started to diminish with increasing time. Finally, Toppino and Bloom suggested that a dual-process model that seeks to account for differences between cued and free recall tasks should not attribute the differences to encoding failures; instead, it should attribute the differences to different retrieval mechanisms that are activated depending upon memory task.

There are three objectives of the present experiments. First, I investigated the importance of both number of intervening items and time between repeated stimuli in generating the spacing and lag effects. Glanzer (1969) demonstrated the importance of number of intervening items in producing the spacing effect, whereas Hintzman et al. (1975b) and Hintzman and Rogers (1973) suggested that time is a key factor. A better understanding of these factors may lead to a revised view of the different theories of encoding.

Second, the present research seeks to better understand under which circumstances spacing and lag effects are found in an incidental recognition task. Research has demonstrated no spacing effect under such conditions (Greene, 1989, 1990; Toppino & Bloom, 2002). However, Block, Bosco, and Schanz (2006) reported that repetition of pictorial stimuli (specifically, human faces) in incidental conditions can lead to a spacing effect. These results are also supported by a recent meta-analysis by Janiszewski et al. (2003). They concluded that the spacing effect is larger under an intentional-memory condition than under incidental-memory conditions, and that it is larger with pictorial stimuli than wit verbal stimuli. Neither Greene nor Toppino and Bloom used pictorial stimuli in their experiments. By using both word stimuli, which can be rehearsed, and pictorial stimuli, which apparently cannot be rehearsed, a better understanding of the role of stimulus type in producing spacing and lag effects can be gained.

Third, the present research investigated the limits of the lag effect. Toppino and Bloom (2002) suggested that lags over 25 s are detrimental to the lag effect. Although Toppino and Bloom measured the inverted-U shape as a function of time, Verkoeijen, Rikers, and Schmidt (2005) demonstrated separate inverted-U shaped relationships as a function of number of intervening items for intentional and incidental conditions. Their findings suggest that performance will increase with up to 14 intervening items in an intentional condition and an increase only up to 8 intervening items in an incidental condition. The current protocol uses large lags—up to 128 items, with an ISI up to 512 s. Using such large lags, I hoped to find a point at which a downturn in recognition performance would be observed. It is thought that considerably larger lags are warranted in recognition memory because performance on recognition memory tasks is higher than performance on free recall tasks. All three objectives seek to paint a comprehensive picture of spacing and lag effects within the framework of recognition memory tasks. This picture includes the role of time and items, the overall shape of these effects, and the peak at which point performance is no longer enhanced with increasing lag.

For the following experiments, it is predicted that significant spacing and lag effects will exist. Thus, it is predicted that immediately repeated items will be more poorly recognized than items repeated with one or more intervening items (the spacing effect). It is also predicted that as the number of intervening items (lag) increases, so too will memory performance as measured by the accuracy of recognition judgments. Specifically, it is predicted that this increase will be observed as a linear trend with a quadratic component. If time is important, there will be no significant differences for total time between presentations across lags (i.e., an ISI of 8 s will yield the same recognition performance regardless of whether 2 items or 4 items fill the ISI). However, if spacing is important, there will be no significant differences for total number of items (i.e., 8 intervening items will yield the same recognition performance regardless of whether the ISI is 16 s or 32 s). Abundant research has repeatedly demonstrated that once-presented items will be remembered worse than repeated items, and also that items benefit from increased study time. It is predicted that repeated items will be recognized significantly better than once-presented items, and also that there will be a significant effect of duration.

EXPERIMENT 1

The objective of Experiment 1 was to study spacing and lag effects using a modified version of Greene's (1989) protocol. A secondary objective was to observe a possible downturn in memory performance as a function of either number of intervening items or ISI between presentation one and presentation two (P1-P2 ISI). The final objective of the current experiment was to determine whether number of intervening items, time between presentations of repeated stimuli, or both are influential.

Method

Participants

A total of 92 undergraduate students enrolled in an introductory psychology course participated in this experiment for course credit. The participants included 58 females and 34 males. The mean age of the participants was 19.8 years.

Design and Materials

A stimulus set was constructed of 180 unique words (see Appendix A) selected from a larger list generated using the English Lexicon Project (Balota et al., in press). The stimuli ranged in Kucera-Francis (1969) frequency ratings from 30 to 597 per million and also ranged in concreteness ratings from 512 to 642. Of the 180 unique words, 100 were study words, 50 were non-studied test words, and 30 were filler words. Four randomized lists were created using the 100 study words and 30 filler words. Filler words were placed in the first and last 15 serial positions of each list to help reduce serial position effects. Some of the filler words were repeated to help eliminate the von Restorff (1933) effect at the beginning and a recency effect at the end of the study list. The 100 study words were composed of 10 subgroups, each containing 10 words. Subgroups were: words presented once, S = 0 (i.e., massed presentation), S = 1, S = 2, S = 4, S = 8, S = 16, S = 32, S = 64, and S = 128, where S represents the total number of intervening items between repetitions. A randomized test list was created using the 100 studied words and the 50 non-studied words.

A pencil and paper survey was constructed containing the PANAS mood inventory (Watson, Clark, & Tellegen, 1998) as a filler task, a 150-item recognition task, and a brief questionnaire. This questionnaire sought to identify participants' gender and age.

Procedure

Stimuli were presented on a large screen using a commercially available projector connected to a PC-compatible computer. E-Prime (Schneider, Eschman, & Zuccolotto, 2002) was used to present the instructions and stimuli.

Participants were tested in groups of no more than six. They were told that they were about to participate in an experiment concerning item judgment and mood, and they consented to participate. They were instructed that they would see a series of words, some of which may be repeated, after which they would be asked about their mood. Next, the participants were shown three sample words similar to the stimuli to which they would be asked to pay attention. Next, one of the four randomized and counterbalanced study lists was presented at either a 2-s or a 4-s rate. In the 2-s condition each word was presented for 1750 ms and was immediately followed by a blank 250 ms inter-stimulus interval (ISI). Each word in the 4-s condition was presented for exactly 3500 ms and was immediately followed by a blank 500 ms ISI.

After the study list was shown, participants were given approximately 2 minutes to describe their current mood using the PANAS mood inventory (Watson et al., 1998). Upon completion of the PANAS, participants were instructed to complete a 150-item recognition memory task. Each item on the recognition task was presented for 4750 ms, and there was a blank 250 ms ISI between test items. Every test item had a sequential number centered above it in parentheses to indicate which test item was currently being shown. Participants were instructed to circle *OLD* on the recognition test if they thought the word was presented earlier during the experiment, and told to circle *NEW* on the test if they thought the word was not presented earlier during the experiment. Finally, the participants were asked to complete a brief demographic questionnaire.

Results

Repeated-measures ANOVAs each using two orthogonal planned comparisons were conducted for the hypothesized main effects of number of intervening items and total inter-stimulus intervals (ISI). These ANOVAs were conducted using both raw hit rates and P_r discrimination indices. Mean hit rates for S = 0, 1, 2, 4, 8, 16, 32, 64, and 128 in the 2-s condition were: .70, .63, .65, .61, .67, .65, .65, .67, and .70 respectively (SE_{pooled} = .02). The mean false-alarm rate in the 2-s condition was .31 (SE = .02). Mean hit rates for these spacings in the 4-s condition were .68, .67, .68, .67, .68, .71, .68, .75, and .72 respectively (SE_{pooled} = .019). The mean false-alarm rate in the 4-s condition was .28 (*SE* = .02).

The number of intervening items had a marginally significant effect on the spacing effect observed with raw hit rate, $F(1, 90) = 3.30$, $p = .07$. The linear trend across lags was significant, $F(1, 90) = 7.16$, $p = .01$, $\eta_p^2 = .07$, but the quadratic trend was not significant $F < 1$. When viewed in terms of ISI rather than spacing, a marginally significant spacing effect was found, $F(1, 90) = 2.96$, $p = .09$, and the linear trend across lags was significant, $F(1, 90) = 6.23$, $p = .01$, $\eta_p^2 = .065$, but the quadratic trend was not significant, $F < 1$.

P_r, a discriminability index that takes into account each individual's hit rate and false alarm rate, was calculated using the formula given by Snodgrass and Corwin (1988)[1]. Figure 1 shows P_r as a function of number of intervening items. Figure 2 shows P_r as a function of ISI. Planned comparisons using P_r revealed a marginally significant spacing effect for number of intervening items, $F(1, 90) = 2.92$, $p = .09$, and ISI, $F(1, 90) = 2.85$, $p = .10$. The linear trend across lags was again observed for both number of intervening items, $F(1, 90) = 10.76$, $p = .001$, $\eta_p^2 = .11$, and ISI, $F(1, 90) = 7.03$, $p = .01$, $\eta_p^2 = .07$. Quadratic trends were not found for items, $F < 1$, or time, $F < 1$. B_r, a measure of bias indicating the probability of saying *OLD* in an uncertain state (Snodgrass & Corwin, 1988), was .46 in both the 2-s and 4-s conditions.

[1] None of these analyses were different using other measures of discriminabilty such as *d′*.

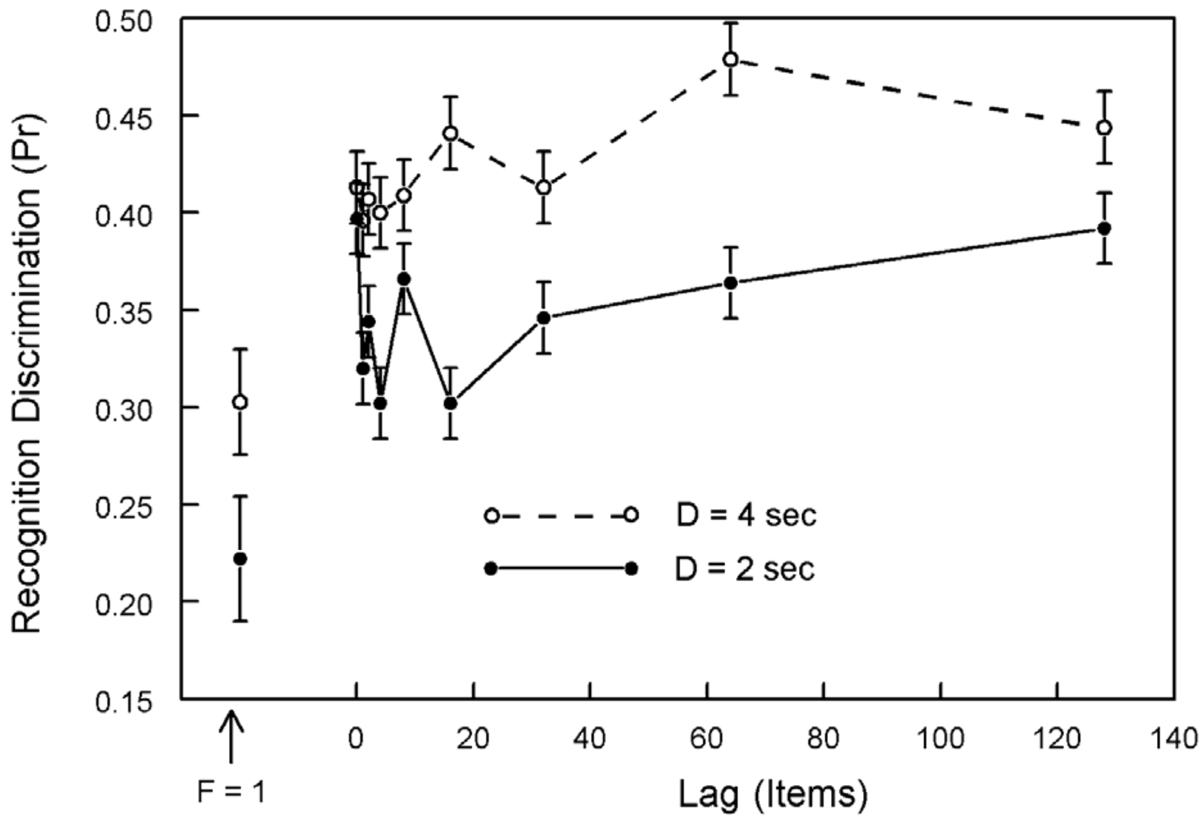

Figure 1. Mean P2 recognition discrimination (P_r) as a function of intervening items between P1 and P2 (lag). Common error bars are based on the within-subjects standard error of the mean (Loftus & Masson, 1994).

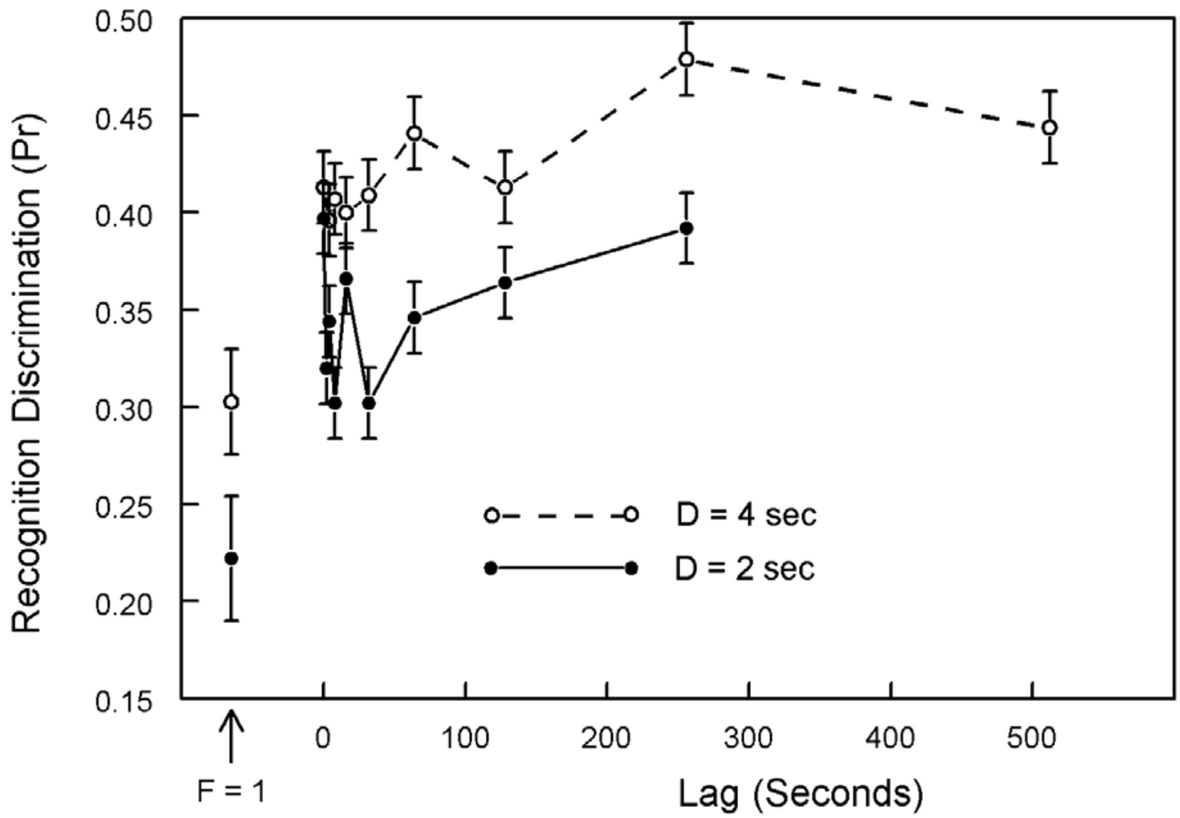

Figure 2. Mean P2 recognition discrimination as a function of P1-P2 ISI. Common error bars are based on within-subjects standard error of the mean (Loftus & Masson, 1994).

It is interesting to note that when the data were analyzed as a function of stimulus duration, both a significant spacing, $F(1, 44) = 4.27$, $p < .05$, $\eta_p^2 = .09$, and lag effect, $F(1, 44) = 6.80$, $p = .01$, $\eta_p^2 = .09$, were found for the 2-s condition ($M = 0.33$, $SE = .02$), whereas only a lag effect, $F(1, 46) = 4.13$, $p < .05$, $\eta_p^2 = .08$, was observed in the

4-s condition (*M* = 0.40, *SE* = .02). Independent-samples *t* tests were conducted to examine the hypothesis that either time or number of intervening items is more important. The comparisons for spacing were grouped by total lag (0, 1, 2, 4, 8, 16, 32, 64, and 128), whereas the comparisons for time were grouped by total ISI (0, 4, 8, 16, 32, 64, 128, and 256 s). For number of intervening items, significant differences between groups were only found in the following comparisons: S = 1 [$t(90) = 2.20$, $p = .03$, $d = 0.46$]; S = 4 [$t(90) = 2.16$, $p = .03$, $d = 0.45$]; S = 16 [$t(90) = 2.39$, $p = .02$, $d = 0.50$]; and S = 64 [$t(90) = 2.59$, $p = .01$, $d = 0.54$]. All other differences were not significant, all *p*s > .17. When comparisons were made by ISI, significant differences existed only in the following comparisons: ISI = 8 [$t(90) = 2.01$, $p < .05$, $d = 0.42$]; ISI = 32 [$t(90) = 1.99$, $p = .05$, $d = 0.41$]; and ISI = 64 [$t(90) = 2.10$, $p = .04$, $d = 0.44$]. All other differences were not significant, all *p*s > .12.

There was a significant difference between items presented once (*M* = .26, *SE* = .034) and repeated twice (*M* = .36, *SE* =.02), $t(91) = 6.40$, $p < .0001$, $d = .60$. A significant effect of stimulus duration was also observed. Words presented at a 4-s rate (*M* =.33, *SE* = .02) were recognized better than words presented at a 2-s rate (*M* = .40, *SE* = .02), $F(1, 90) = 4.67$, $p = .033$, $d = 0.45$. B_r, a measure of bias indicating the probability of saying *OLD* in an uncertain state (Snodgrass & Corwin, 1988), was .46 in both the 2-s and 4-s conditions.

Discussion

The results of Experiment 1 show only partial agreement with the findings of previous research including Greene (1989) and Toppino and Bloom (2003). An incidental memory paradigm using words, much like the procedure used by Greene, produced a spacing effect when the stimuli were presented at a presentation rate of 2 s, but not 4 s. Significant lag effects were also observed using large lags.

Two unusual findings can be seen in these data. First, the significant spacing effect in the 2-s condition was in the opposite direction from commonly found spacing effects: Recognition of S = 0 words was better than recognition of words with S > 0. A possible explanation for this, as discussed further in the general discussion, is the von Restorff (1933) effect. Second, while significant lag effects were observed for both conditions, they did not follow the predicted pattern, nor did they follow other patterns commonly reported in the spacing effect literature: They only showed a linear trend, whereas most previous studies have also found a quadratic trend. This suggests that there may be no observable downturn in memory performance at long lags. These findings are unique, and if replicated using stimuli that are not able to be rehearsed, would be important findings in the spacing effect literature.

Analyses conducted to determine whether time or total number of intervening items between repeated items contribute more heavily to the spacing effect yielded an interesting finding. One would predict that if either factor were important, significant differences would not be seen in separate *t* tests for the separate levels of each factor.

However, significant differences were found in separate comparisons for both items and time. Although these differences were not predicted, the finding that differences existed for both time and items could suggest that both are important factors in generating the spacing effect. A goal of the second experiment is to add to the previous data to further evaluate the importance of each of these factors.

EXPERIMENT 2

The objectives of Experiment 2 were identical to the objectives listed for Experiment 1. The procedural difference of using stimuli that theoretically cannot be rehearsed allows for a direct test of the theory that rehearsal affects recognition. Although we could assume no maintenance rehearsal was occurring, elaborative rehearsal could have been occurring. It was expected that the results of Experiment 2 would be similar to those observed in Experiment 1.

Method

Participants

A total of 60 undergraduate students enrolled in an introductory psychology course participated in this experiment for course credit. The participants included 35 females and 24 males. The mean age of the participants was 21.4 years.

Design and Materials

A stimulus set was constructed of 170 unique faces selected from a larger set of faces used by Carmel and Bentin (2002)[2]. The pictures used showed individual's heads and shoulders with neutral facial expressions and were photographed in black and white. Sample pictures are shown in Appendix B. The stimulus set contained 84 female and 86 male faces. Of the 170 unique faces, 100 were study items, 50 were non-studied test

---

[2] I would like to thank Shlomo Bentin for providing Richard Block with the face stimuli used in this experiment.

items, and 20 were filler items. Three randomized presentation sets and a test set were generated using the same methods described in Experiment 1.

The same pencil and paper survey as Experiment 1 containing the PANAS mood inventory (Watson et al., 1998) as a filler task, a 150-item recognition task, and a brief questionnaire was used.

Procedure

The procedure for Experiment 2 was exactly the same as Experiment 1, with the following exception: During the test phase, participants saw a number in parentheses to indicate the test item they were about to see for a duration of 1000 ms. The test item then followed for a duration of 4000 ms. This change was added to facilitate the proper alignment of each test item on the computer screen due to the programming constraints of E-Prime (Schneider et al., 2002).

Results

Repeated-measures ANOVAs each using two orthogonal planned comparisons were conducted for the hypothesized main effects of number of intervening items and total ISI. These ANOVAs were conducted using both raw hit rates and also P_r discrimination indices. Mean hit rates for S = 0, 1, 2, 4, 8, 16, 32, 64, and 128 in the 2-s condition were .63, .59, .57, .65, .62, .61, .72, .62, and .73 respectively (SE_{pooled} = .03). In this condition, the mean false-alarm rate was found to be .28 (*SE* = .03). Mean hit rates for these spacings in the 4-s condition were . 70, .67, .67, .68, .68, .66, .70, .64, and .71

respectively (SE_{pooled} = .02). In the 4-s condition, the mean false-alarm rate was .29 (*SE* = .03).

The number of intervening items had no significant effect on the spacing effect observed with raw hit rate, $F < 1$. However, the linear trend across lags was observed, $F(1, 59) = 12.68$, $p < .0001$, $\eta_p^2 = .18$. No quadratic trend across lags was observed, $F < 1$. When viewed in terms of ISI rather than spacing, no significant spacing effect was found $F < 1$, but a significant linear trend was observed across lags, $F(1, 59) = 6.71$, $p = .01$, $\eta_p^2 = .10$. Again, the quadratic trend was not significant, $F < 1$.

Figure 3 shows P_r as a function of number of intervening items. Figure 4 shows P_r as a function of ISI. Planned comparisons using P_r revealed no significant spacing effect for either number of intervening items, $F < 1$, or inter-stimulus interval $F < 1$. The linear trend across lags was again significant for both number of intervening items, $F(1, 59) = 12.55$, $p = .001$, $\eta_p^2 = .07$, and ISI, $F(1, 59) = 4.35$, $p = .04$, $\eta_p^2 = .07$. The quadratic trend was not significant for either items, $F < 1$, or time, $F < 1$. B_r was .43 in the 2-s condition and .47 in the 4-s condition.

When analyzed by duration, a significant linear trend across lags was only found in the 2-s condition. This trend was found for both number of intervening items, $F(1, 30) = 12.00$, $p = .001$, and ISI, $F(1, 30) = 11.84$, $p = .002$. This effect was not significant for both items, $F(1, 29) = 1.50$, $p = .23$, and ISI, $\mathrm{F}(1, 29) = 1.50$, $p = .23$, in the 4-s condition.

As in Experiment 1, independent-samples *t* tests were conducted to analyze whether number of items or time contributed more to the spacing effect. When analyzed for number of intervening items, no comparison was significant (all *p*s > .12). Analyzed for ISI, only a marginally significant difference was found in the 256-s comparison $t(59) = 1.75$, $p = .086$, $d = 0.44$, with all other comparisons not significant, all *p*s > .15.

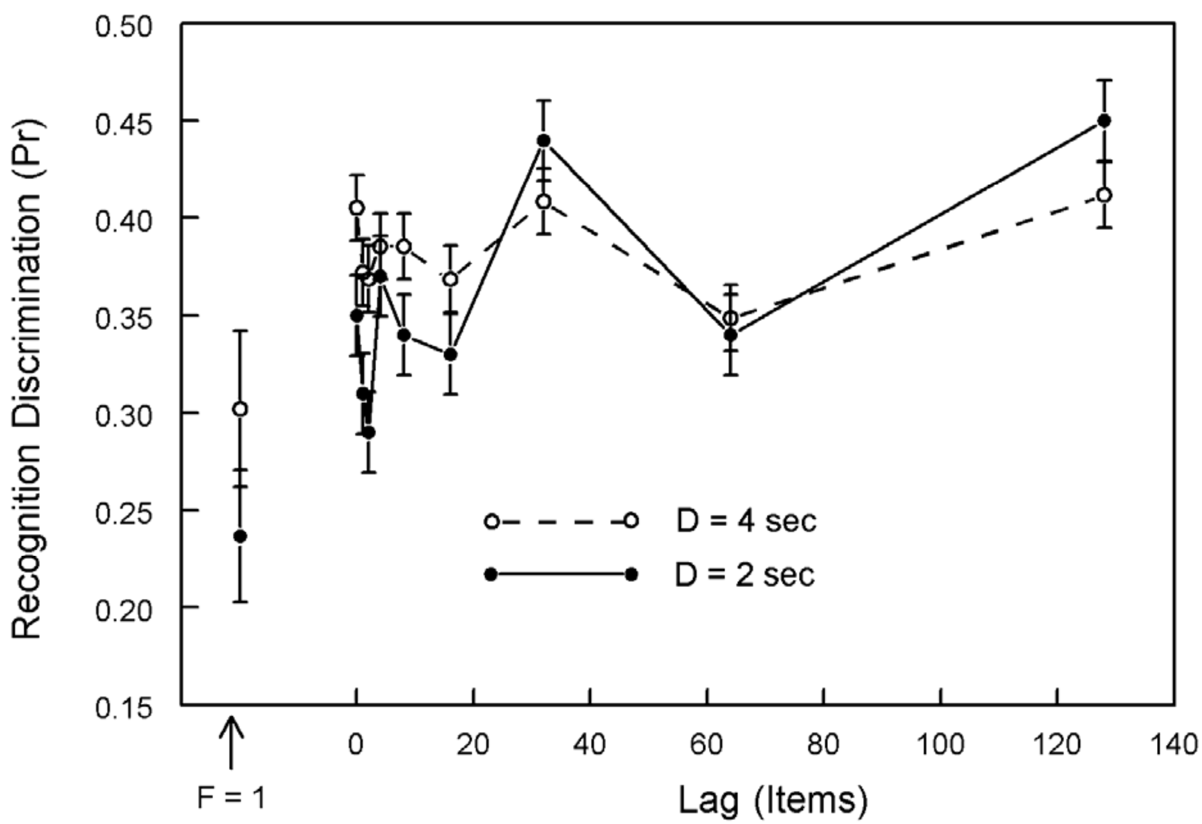

Figure 3. Mean P2 recognition discrimination (P_r) as a function of intervening items between P1 and P2 (lag). Common error bars are based on the within-subjects standard error of the mean (Loftus & Masson, 1994).

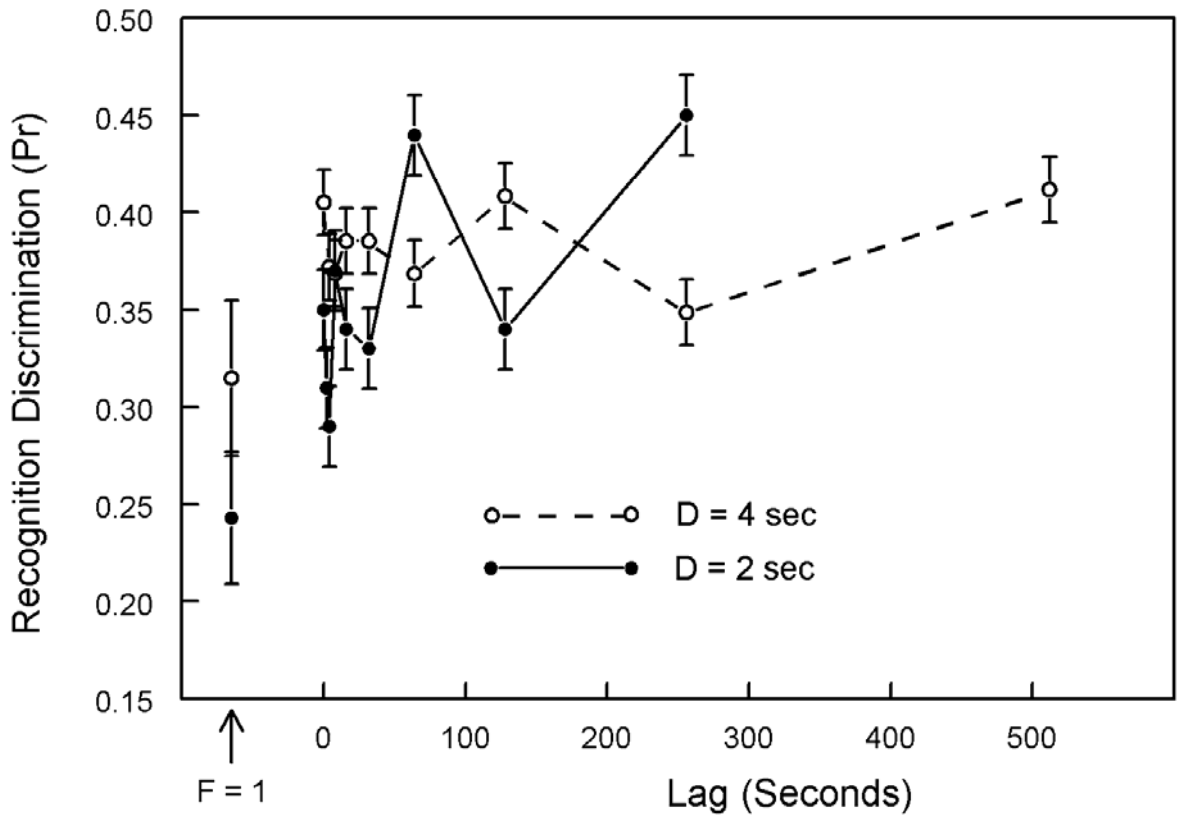

Figure 4. Mean P2 recognition discrimination (P_r) as a function of P1-P2 ISI. Common error bars are based on the within-subjects standard error of the mean (Loftus & Masson, 1994).

Once presented items (*M* = 0.27, *SE* = .04) were recognized less than repeated items (*M* = 0.35, *SE* = .02), $t(60) = 3.49$, $p = .001$, $d = 0.89$. Unlike Experiment 1, no significant effect of stimulus duration was observed, $F < 1$.

Discussion

The results for Experiment 2 show a contrast to the results obtained in Experiment 1. Perhaps the most interesting finding to highlight is the apparent lack of an effect of duration. Two possible explanations exist for this finding. First, faces are known to be encoded rapidly. Second, faces should not benefit from rehearsal. Taken together, this could suggest that 2-s presentations of face stimuli are sufficient to produce the same level of familiarity needed for recognition as a 4-s stimulus.

Although a lag effect was observed in both conditions in Experiment 1, a significant lag effect was only found for the 2-s condition in Experiment 2. In contrast to Experiment 1, no spacing effects were observed. One similarity between the observations in Experiment 1 and Experiment 2 is the lack of a downturn in memory performance. Because the data of neither experiment showed a downturn even though long lags were used, previous findings that suggest the shape of the spacing and lag effects are an inverted-U are called into question. As in Experiment 1, it is impossible from the observed data to separate the effects of time and items in generating the observed lag effect. This finding adds support to the theory that both time and items are important in generating spacing and lag effects.

GENERAL DISCUSSION

The present experiments sought to investigate three important questions regarding spacing and lag effects on recognition-memory performance under incidental-memory conditions. One goal was to observe the effect of stimulus type in the spacing effect. A second goal was to investigate a possible downturn in memory performance at relatively large lags. The final goal of the present experiments was to investigate the importance of time between presentations and number of intervening items between presentations of a repeated item in producing spacing and lag effects. The results of the experiments yielded rather unusual findings.

Perhaps the most interesting finding in the previous experiments is that of a marginally significant spacing effect for words and the absence of a significant spacing effect for faces. The finding of a spacing effect for words is important in that it failed to replicate previous work by Greene (1989, 1990), who also used incidental-memory conditions. In addition, the lack of a spacing effect for faces is unusual because it failed to replicate the findings of others (e.g., Block et al., 2006; Paivio, 1974).

Surprisingly, the spacing effect found in Experiment 1 was not in the predicted direction. Words repeated immediately were recognized better than words repeated with up to four intervening items. When more than four items separated repeated words, there was no difference between massed ($S = 0$) presentations and spaced ($S > 0$) presentations. This finding, added to the linear lag effect without a quadratic component, suggests that under these circumstances, there was a uniquely shaped spacing effect resembling a checkmark. Such an effect has not been previously published in the spacing effect literature. An ad hoc explanation for the increased recognition of massed items could be the existence of a von Restorff (1933) effect. Spacing effect experiments traditionally use a limited set of lags in which massed repetitions account for around 20% of the total trials in the experiment. The proportion of massed items in the present experiments was only about 10%. It is possible that presentations of massed items were so rare that participants paid extra attention to them during the second presentation. This extra attention may have enhanced recognition during the recognition test.

In Experiment 2, in contrast to Experiment 1, there were no significant effects of spacing. Just as it was curious to find this uniquely shaped spacing effect, it appears equally curious not to have found any spacing effect. While these results are similar to those observed by Greene (1989), it is hard to attribute them to the same process, specifically deficient-processing theory, as he originally had. He reasoned that stimuli which appeared familiar would not receive the same amount of rehearsal as those which did not seem familiar. In this instance, it is theoretically impossible for one to rehearse items, whether they were familiar or not. Moreover, these results conflict with those of Block et al. (2006) by not finding any significant spacing effect with pictorial stimuli.

Of considerable importance to the findings concerning the spacing effect is the observation in both experiments of significant effects of lag. While the effect of lag was significant in both experiments, the predicted shape of the effect was not found. Traditionally, the lag effect has been found to be monotonically increasing, yet decelerating. This would be reflected in the finding of both linear and quadratic trends.

Only significant linear trends were observed in the present experiments. The importance of finding only linear trends for lag effects is that there was no observable downturn, or point at which memory performance started to decrease, as a result of large lags. These findings contradict previous research by Toppino and Bloom (2002) and Verkoeijen et al. (2005), which revealed an inverted-U shaped function. It is important, however, to note that the studies which found an inverted-U shaped function also used free recall paradigms, and this may be an important difference to examine further.

In both experiments, the role of time and the role of number of intervening items were both important in creating spacing and lag effects. This finding implies that to sufficiently describe spacing and lag effects a theory must be able to account for the effects of both time and number of intervening items. At least one such theory that may be suited to deal with the role of time and number of items is the encoding variability theory. This theory takes into account that as items increase between P1-P2 repetitions, so too does the number of possible contextual cues that get encoded with the to-be-recognized item. Recursive reminding, a subset of the encoding variability theory, also suggests that time between repetitions is an important contextual cue. Of Hintzman's (1974) original automatic process theories, it is the one that can explain the current results the best, and is also favored by current meta-analyses (Cepeda et al., 2006; Janiszewski et al., 2003).

There are several limitations to the current experiments. The current study used only incidental instructions, limiting the ability to generalize results to the effects of spacing and lag using free recall conditions. Another possible limitation is inherent in the design. As an alternative to hypothesizing that both time and items are important in generating the spacing effect, one could argue that the current experiment was not sensitive enough to detect such differences for two reasons. The first such reason is the inherent confound by manipulating time between subjects. Although the total ISI or number of items varied equally between the 2-s and 4-s conditions, there is still the fact that one experimental group of a comparison viewed the stimuli at a 2-s rate while the other group viewed stimuli at a 4-s rate. A second source of problem in trying to answer this question is that although one can always vary time independently of items, to increase the number of items, one must inherently increase the amount of time between P1-P2 presentations. An alternative that may solve the first problem would be to create a paradigm in which duration, spacing, and ISI are all manipulated using a within-subjects design.

Future Research

Many questions about the spacing effect that remain unanswered. Future research should focus on trying to isolate the conditions under which certain automatic and voluntary processes moderate spacing and lag effects. Within this realm, research should focus on the complex interactions of intentionality and stimulus type. Although the design of this study sought to compare time and items using a between-subjects design, it would be interesting to make the same comparisons using a within-subjects design, thus potentially reducing the confound of stimulus duration. Another interesting research question that arose from the present study is the degree to which instructions within an incidental condition are truly incidental. For example, using words as stimuli and telling participants to decipher a complex rule about a study list (as Greene, 1989, did) inherently requires semantic processing. However, telling participants that their mood will be tested at a later time does not require semantic processing. Understanding the different levels of processing individuals use to encode information is valuable knowledge when looking at the underlying mechanisms of the spacing effect. Finally, one should test the hypothesis that enhanced recognition of massed items was due to a von Restorff effect by using the same instructions and decreasing the total number of conditions such that the overall proportion of massed trials is similar to proportions used in the previous spacing effect literature.

Conclusion

The specific goals of this study were to understand the role of stimulus type in the spacing and lag effects, to find a possible downturn in the lag effect, and to gain a better understanding of the importance of time and number of intervening items in creating the spacing and lag effects. The results obtained were unable to replicate earlier findings for both word stimuli and pictorial stimuli. In fact, opposite results were found. Greene (1989) found no spacing effect using word stimuli in an incidental condition; however, a unique spacing effect in the direction opposite of that predicted was found in Experiment 1. Block et al. (2006) found a spacing effect for faces, along with a lag effect. Experiment 2 found a significant lag effect, but no spacing effect. The significant lag effects found in both experiments showed no clear downturn at large spacings; however, the general shape of the lag effect found in this study does not match a commonly accepted finding that is a monotonically increasing, yet decelerating function of lag. Finally, the results suggested that the effects of time and items cannot be decoupled. Analyses revealed no significant differences in recognition performance as a function of spacing, and also as a function of number of intervening items. The importance of this finding is that it suggests both time and items are important factors contributing to the spacing effect.

With regards to classic one-process and recent dual-process theories of the spacing and lag effect, the experimental results of this study do not seem to fit easily into several of the previous theoretical models. Because memory was better for massed presentations than short spacings (Experiment 1), and was not significantly worse for massed presentations than spaced repetitions (Experiment 2), one could argue that the findings do not support habituation-recovery or consolidation theories. The same findings also do not support a voluntary rehearsal theory because it is implied that memory is worse for massed repetitions because individuals do not rehearse familiar, or already activated, stimuli as much as stimuli that seem unfamiliar. Voluntary attention may play a role in the current experiments because with such a large stimulus list individuals may become bored and not be attending to the stimuli. This could also explain differences between this study and previous studies in which participants had to actively process stimulus information at a semantic level for each item. When participants are engaged in the experiment, they may be less likely to be bored, or let their attention wander. Encoding variability theory, and a sub-theory known as *recursive reminding* may still be able to explain the effects of spacing and lag seen here because it relies on contextual cues that are both external (i.e., semantic and facial features) and internal (i.e., temporal cues).

Based on previous information, it may be a good idea to create a model of the spacing effect that is not dependent on only two processes. It may be useful to create a bottom-up model of the processes that contribute to these effects. Such a model should start by taking into account which mechanisms are online at the start of a learning trial, which could be mediated by intentionality. The next factors that should be taken into account are moderators of encoding. Such moderators could include stimulus type ranging from novel (i.e., nonsense syllables and complex pictures) to highly familiar (i.e., words). Another strong moderator of encoding should be how rehearsable each stimulus is. This can range from theoretically impossible to rehearse (i.e., pictures) to very rehearsable (i.e., words). The next important step to consider in such a model would be the processes that occur at retrieval. Retrieval is most directly affected by retrieval task such as free recall and cued-recall. Even cued-recall tasks have different levels of difficulty depending upon the type of cue used. Each of these levels should be used to form predictions of spacing and lag effects specific to these factors. A model taking into consideration all of these variables is inherently complex and may prove difficult to construct.

Spacing and lag effects have historically been well-studied effects, yet they are not easily understood. In a memory task, there are several points in the timeline of events associated with studying these effects where something can go right or wrong. Each of these points can be influenced by multiple factors. The current theories of spacing and lag tend to focus more heavily on a particular point in the overall memory process rather than having an expanded view of the entire process. A comprehensive model that takes into account the several mediators and moderators of these effects is needed.

REFERENCES

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K.A., Kessler, B., Loftus, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (in press). The English lexicon project: A user's guide. *Behavior Research Methods*.

Block, R. A., Bosco, F. A., & Schanz, T. S. (2006, April). *The spacing effect: Evidence from duration productions and recognition memory*. Poster presented at the meeting of the Western Psychological Association, Palm Springs, CA.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*, 1–29.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.

Glanzer, M. (1969). Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning & Verbal Behavior 8*, 105–111.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 371–377.

Greene, R. L. (1990). Spacing effects on implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1004–1011.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77–99). Hillsdale, NJ: Erlbaum.

Hintzman, D. L. (2004a). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, *32*, 336–350.

Hintzman, D. L. (2004b). Time versus items in judgment of recency. *Memory & Cognition*, *32*, 1298–1304.

Hintzman, D. L., & Rogers, M. K. (1973). Spacing effects in picture memory. *Memory & Cognition, 1*, 430–434.

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975a). Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 31–40.

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975b). What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures. *Memory & Cognition, 3*, 287–294.

Hintzman, D. L., Summers, J. J., Eki, N. T., & Moore, M. D. (1975). Voluntary attention and the spacing effect. *Memory & Cognition, 3*, 576–580.

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research, 30*, 138–149.

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review, 12*, 159–164.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Paivio, A. (1974). Spacing of repetitions in the incidental and intentional free recall of pictures and words. *Journal of Verbal Learning & Verbal Behavior, 13*, 497–511.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 1.1). Pittsburgh, PA: Psychology Software Tools.

Shaffer, W., & Shiffrin, R. M. (1972). Rehearsal and storage of visual information. *Journal of Experimental Psychology*, *92*, 292–296.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50.

Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 437–444.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2005). Limitations to the spacing effect: Demonstration of an inverted u-shaped relationship between interrepetition spacing and free recall. *Experimental Psychology, 52*, 257–263.

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Sprurenfeld [The effects of field formation in the trace field]. *Psychologie Forschung, 18,* 299–334.

Watson, D., Clark, L. A., & Tellegen, A. (1998). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychologly, 54*, 1063–1070.

Waugh, N. C., & Norman, D. A. (1968). The measure of interference in primary memory. *Journal of Verbal Learning & Verbal Behavior, 7*, 617–626.

APPENDICES

APPENDIX A

LIST OF WORD STIMULI USED IN EXPERIMENT 1

WORD LIST FOR EXPERIMENT 1

animal
army
artist
avenue
baby
ball
band
bank
battle
beach
bear
beef
bench
bill
bird
block
blood
board
boat
body
bomb
bone
book
bottle
brain
branch
bread
bridge
brush
camera
camp
case
cash
cattle
cell
cent
chain
chair
chest
child
circle
city
clay
cloth
coal
coast
coat
coffee
colors
common
corn
corner
crew
cross
crowd
damage
date
desk
dinner
dirt
doctor
dollar
door
dress
drill
drink
driver
dust
earth
engine
estate
face
family
farm
feet
fence
film
fire
fish
flesh
floor
foam
food
foot
forest
fort
frame
fruit
garden
gate
gift
girl
glass
gold
golf
grass
grave
ground
guard
guest
hair
hall
hand
head
heart
hill
home
horn
horse
hotel
house
human
iron
island
jacket
jury
king
knee
knife
lady

lake
land
lane
lawyer
lead
letter
light
liquid
lumber
lunch
male
mantle
market
match
meal
meat
metal
money
moon
mother
motor
mouth
muscle
music
neck
nose
note
novel
ocean
office
page
paint
palace
paper
park
pencil
person
phone
piano
plane
plant
pocket
poet
pool
porch
post
prince
prison
queen
radio
rain
record
rector
rice
rifle
ring
river
road
rock
roof
room
root
rose
salt
school
seat
seed
sheet
ship
shop
shore
sign
sister
skin
smile
smoke
snake
snow
soil
spoke
spring
square
stable
staff
steel
stick
stone
store
street
sugar
suit
supper
table
tape
teeth
temple
test
throat
tool
town
track
train
tree
truck
tube
uncle
valley
wall
water
weapon
wheel
wife
wind
window
wine
wire
wood
worker
yard
yellow

APPENDIX B

SAMPLE FACE STIMULI USED IN EXPERIMENT 2

SAMPLE FACE STIMULI FOR EXPERIMENT 2