INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
RICE UNIVERSITY

THE SPACING EFFECT: IMPLICATIONS FOR RELEARNING

by

DANA M. SCHNEIDER

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE

David M. Lane, Director
Associate Professor of Psychology

John W. Brelsford
Professor of Psychology

Miguel A. Quiñones
Assistant Professor of Psychology

Eva J. Lee
Professor of Human Performance
and Health Science

Houston, Texas
May 1997
ABSTRACT

The Spacing Effect: Implications for Relearning

by

Dana M. Schneider

Experiments reported here examined the effects of distributing practice during relearning. Specifically, they provide an initial test of the prediction by R. A. Bjork and C. O. Fritz (1994), based on the new theory of disuse (R. A. Bjork & E. L. Bjork, 1992), that spacing practice is not important for relearning. In Experiment 1a, the speed to respond to simple numeric multiplication problems was measured after subjects practiced the problems under three different relearning schedules: (1) A massed condition in which all of the practice on a specific problem occurred consecutively, (2) a spaced condition in which there was a uniform spacing of one intervening problem between each practice on a particular problem, and (3) an expanded condition in which the practice was spaced in an expanded fashion, such that first there were no intervening problems between practices, then there were 4 problems, followed by 8. No significant differences among these three conditions were found. In Experiment 1b, original learning on an analogous task, mental arithmetic involving letters rather than numbers, was performed under the same three learning schedules. A spacing effect was found under
the original learning conditions of Experiment 1b. Experiment 2 was similar to Experiment 1a, except that the instructions placed more stress on accuracy in order to avoid problems associated with a speed/accuracy trade-off. As with Experiment 1a, there was no significant effect of condition. The results provide preliminary support for Bjork and Fritz's prediction regarding the distribution of practice during relearning. The findings are encouraging from a practical standpoint and support the notion that level of expertise is an important factor in research on long-term retention.
ACKNOWLEDGMENTS

I would like to express my sincere appreciation to everyone who helped make my dream of completing my dissertation a reality. First, I would like to thank my advisor, David Lane, for his guidance and encouragement throughout this project. I would also like to thank my other committee members: John Brelsford, Miguel Quiñones, and Eva Lee. I am grateful to all of them for helping me succeed in achieving my goal. I would also like to give special thanks to Robert Bjork for being a constant mentor and friend. He will always be like a father to me.

Amongst the other people who directly contributed in making this dissertation possible, I would like to thank Lance Bloom, whose Hypercard expertise provided me with invaluable help in programming the experiments. In addition, I would like to thank Angie Yanez for being such a fabulous research assistant—not only on this project, but throughout my time at Rice University. I would also like to thank the rest of my class, Roman Longoria and Justin Miller, for their critical comments on previous drafts of this dissertation.

I am especially grateful for the constant support and unyielding confidence in me from all of my friends and family. In particular, I would like to send my love and thanks to my classmate, friend, and soon-to-be husband, Justin Miller. Most of all, I would like to thank my mom. Words cannot portray just how grateful I am to her and how much I love her. Finally, I wish to dedicate this dissertation to my grandfather, Earl Fifer (a.k.a., "Papa"). My thoughts are with him.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background on the Spacing Effect</td>
<td>2</td>
</tr>
<tr>
<td>Generality of the Spacing Effect</td>
<td>6</td>
</tr>
<tr>
<td>Verbal Learning and Memory</td>
<td>6</td>
</tr>
<tr>
<td>Motor Skill Acquisition</td>
<td>8</td>
</tr>
<tr>
<td>&quot;Test-Spacing Effect&quot;</td>
<td>9</td>
</tr>
<tr>
<td>Current Theories</td>
<td>11</td>
</tr>
<tr>
<td>Deficient Processing Theories</td>
<td>13</td>
</tr>
<tr>
<td>Encoding Variability Theories</td>
<td>14</td>
</tr>
<tr>
<td>The New Theory of Disuse</td>
<td>15</td>
</tr>
<tr>
<td>Implications for the Spacing Effect</td>
<td>17</td>
</tr>
<tr>
<td>Rationale for the Present Research</td>
<td>22</td>
</tr>
<tr>
<td>Considerations for Selecting the Task and</td>
<td>26</td>
</tr>
<tr>
<td>Design of the Present Research</td>
<td></td>
</tr>
<tr>
<td>Pilot Research</td>
<td>27</td>
</tr>
<tr>
<td>Present Research</td>
<td>29</td>
</tr>
<tr>
<td>Experiment 1a</td>
<td>30</td>
</tr>
<tr>
<td>Method</td>
<td>32</td>
</tr>
<tr>
<td>Subjects</td>
<td>32</td>
</tr>
<tr>
<td>Design</td>
<td>32</td>
</tr>
<tr>
<td>Materials</td>
<td>32</td>
</tr>
<tr>
<td>Study Lists</td>
<td>33</td>
</tr>
<tr>
<td>Retention Test</td>
<td>37</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distractor Task</td>
<td>37</td>
</tr>
<tr>
<td>Procedure</td>
<td>40</td>
</tr>
<tr>
<td>Results</td>
<td>41</td>
</tr>
<tr>
<td>Acquisition Test Trials</td>
<td>41</td>
</tr>
<tr>
<td>Accuracy</td>
<td>41</td>
</tr>
<tr>
<td>Reaction Times</td>
<td>41</td>
</tr>
<tr>
<td>Retention Test</td>
<td>45</td>
</tr>
<tr>
<td>Accuracy</td>
<td>45</td>
</tr>
<tr>
<td>Reaction Times</td>
<td>45</td>
</tr>
<tr>
<td>Wonderlic Personnel Test</td>
<td>47</td>
</tr>
<tr>
<td>Discussion</td>
<td>47</td>
</tr>
<tr>
<td>Experiment 1b</td>
<td>49</td>
</tr>
<tr>
<td>Method</td>
<td>50</td>
</tr>
<tr>
<td>Subjects</td>
<td>50</td>
</tr>
<tr>
<td>Design</td>
<td>50</td>
</tr>
<tr>
<td>Materials</td>
<td>51</td>
</tr>
<tr>
<td>Study Lists</td>
<td>51</td>
</tr>
<tr>
<td>Retention Test</td>
<td>54</td>
</tr>
<tr>
<td>Distractor Task</td>
<td>55</td>
</tr>
<tr>
<td>Procedure</td>
<td>55</td>
</tr>
<tr>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>Acquisition Test Trials</td>
<td>57</td>
</tr>
<tr>
<td>Retention Test</td>
<td>57</td>
</tr>
<tr>
<td>TABLE OF CONTENTS (continued)</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Wonderlic Personnel Test</strong></td>
<td>Page</td>
</tr>
<tr>
<td>Discussion</td>
<td>60</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>64</td>
</tr>
<tr>
<td>Method</td>
<td>65</td>
</tr>
<tr>
<td>Subjects</td>
<td>65</td>
</tr>
<tr>
<td>Design</td>
<td>65</td>
</tr>
<tr>
<td>Materials</td>
<td>65</td>
</tr>
<tr>
<td>Procedure</td>
<td>65</td>
</tr>
<tr>
<td>Results</td>
<td>65</td>
</tr>
<tr>
<td>Acquisition Test Trials</td>
<td>66</td>
</tr>
<tr>
<td>Accuracy</td>
<td>66</td>
</tr>
<tr>
<td>Reaction Times</td>
<td>66</td>
</tr>
<tr>
<td>Retention Test</td>
<td>69</td>
</tr>
<tr>
<td>Accuracy</td>
<td>69</td>
</tr>
<tr>
<td>Reaction Times</td>
<td>69</td>
</tr>
<tr>
<td>Wonderlic Personnel Test</td>
<td>71</td>
</tr>
<tr>
<td>Discussion</td>
<td>71</td>
</tr>
<tr>
<td>General Discussion</td>
<td>73</td>
</tr>
<tr>
<td>Theoretical Implications</td>
<td>73</td>
</tr>
<tr>
<td>Practical Implications</td>
<td>74</td>
</tr>
<tr>
<td>Benefits of Massed Practice</td>
<td>74</td>
</tr>
<tr>
<td>Subjective Preference for Massed Practice</td>
<td>75</td>
</tr>
<tr>
<td>Scheduling Benefits of Massed Practice</td>
<td>76</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Learning vs. Relearning</td>
<td>77</td>
</tr>
<tr>
<td>Related Research</td>
<td>79</td>
</tr>
<tr>
<td>Possibilities for Future Research</td>
<td>81</td>
</tr>
<tr>
<td>References</td>
<td>84</td>
</tr>
<tr>
<td>Appendix</td>
<td>94</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Critical multiplication problems for Experiments 1a and 2 ......................................................... 34
Table 2. Filler multiplication problems for Experiments 1a and 2 ................................................................. 35
Table 3. Counterbalancing of multiplication problems in Experiments 1a, 1b, and 2 ............................................. 36
Table 4. Sequencing of Massed (M), Spaced (S), Expanded (E), and Filler (F) trials for study lists in Experiments 1a, 1b, and 2 .................................................................................. 38
Table 5. Median serial position in the study lists for the experimental conditions ......................................................... 39
Table 6. Intercorrelations between Wonderlic Personnel Test scores and median correct reaction times on the retention test in Experiment 1a ............................................................................. 48
Table 7. Critical multiplication problems for Experiment 1b ................................................................................. 52
Table 8. Filler Multiplication Problems for Experiment 1b ...................................................................................... 53
Table 9. Intercorrelations between Wonderlic Personnel Test scores and proportion of correct responses on the retention test in Experiment 1b .............................................................................. 61
Table 10. Intercorrelations between Wonderlic Personnel Test scores and median correct reaction times on the retention test in Experiment 2 ............................................................................ 72
LIST OF FIGURES

Figure 1. Temporal sequence of presentations and test under typical massed and spaced conditions.......................12

Figure 2. Retrieval strength (R) and storage strength (S) as a function of practice under massed original learning conditions..............................................................................19

Figure 3. Retrieval strength (R) and storage strength (S) as a function of practice under spaced original learning conditions..............................................................................20

Figure 4. Retrieval strength / performance for massed and spaced conditions during the acquisition trials and on the retention test during original learning..............................21

Figure 5. Retrieval strength (R) and storage strength (S) as a function of practice under massed relearning conditions..................................................................................23

Figure 6. Retrieval strength (R) and storage strength (S) as a function of practice under spaced relearning conditions..................................................................................24

Figure 7. Retrieval strength / performance for massed and spaced conditions during the acquisition trials and on the retention test during relearning..............................25
LIST OF FIGURES (continued)

Figure 8. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1a.................................42

Figure 9. Average median reaction times (in sec) of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1a.................................44

Figures 10a and 10b.
Boxplots of the mean and median correct reaction times (in sec) for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 1a......................46

Figure 11. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1b.................................58

Figure 12. Boxplots of the proportion of correct responses for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 1b............................59

Figure 13. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 2.................................67
LIST OF FIGURES (continued)

Figure 14. Average median reaction times (in sec) of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 2........................................68

Figures 15a and 15b.
Boxplots of the mean and median correct reaction times (in sec) for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 2.........................70
The Spacing Effect: Implications for Relearning

Throughout our lives we are confronted with the task of acquiring information and skills in a wide array of domains. Whether it be learning information in school, on-the-job training for one's occupation, or motor-skill acquisition for a particular sport, a common question arises: What is the best way to master the information or skill? This question, one that has applicability in numerous aspects of our lives, has attracted much attention in the field of experimental psychology.

The well-known phrase "practice makes perfect" keys in on the fact that repetition or rehearsal is important for learning. However, not all practice is equally effective. For example, if one is trying to learn a list of foreign vocabulary words, repeatedly studying one word at a time before moving on to the next word in the list is not as effective as studying each word for a short time on a number of different instances (i.e., study on a particular word is interleaved with study on other words). Similarly, if one is trying to learn how to shoot a free-throw in basketball, practicing a couple of hours on two different days will most likely lead to greater success (in the long-term) than practicing four hours all on the same day. These examples are generalizations of a phenomenon typically referred to as the spacing effect.

Specifically, the spacing effect is that, for a fixed amount of study time, spaced practice is better for long-term retention than is massed practice. Practice is considered to be spaced (or distributed) if the practice episodes are separated in time, rather than being massed (or contiguous) in time.
Background on the Spacing Effect

The spacing effect was one of the first research topics investigated by experimental psychologists. For example, Ebbinghaus (1885/1964) concluded that "with any considerable number of repetitions a suitable distribution of them over a space of time is decidedly more advantageous than the massing of them at a single time" (p. 89). Ebbinghaus' finding was formally stated in the form of Jost's Law in 1897 which states, "If two associations are of equal strength but of different age, a new repetition has a greater value for the older one." (McGeoch, 1943, p. 140)

A typical experiment on the spacing effect begins with an acquisition or training phase. (Note: The term 'training' is defined as acquisition of the crucial skills or information to perform a given task, see Druckman & Bjork, 1991.) During acquisition, the learner, or subject, is presented with information to study or given practice on a given task under massed and/or spaced conditions. Although a measure of performance can be obtained during the acquisition or training phase, the experimenter is typically more interested in the learner's performance at some later time. Temporary effects from the experimental manipulations, such as increased performance from motivational factors or decreased performance from fatigue, may have large effects on the level of performance during acquisition. Consequently, performance during acquisition is usually a poor indicator of the relative amounts of learning and understanding in the different experimental conditions (see Druckman & Bjork, 1991; Schmidt & Bjork, 1992). In order for one experimental condition to lead to more learning than another, differences in performance must have "some
permanence across time, or...be able to survive the removal of the manipulation in question" (Schmidt & Bjork, 1992, p. 208). In most educational, sports, military, and industrial settings, it is performance in the long-term that is of utmost importance (see Bjork, 1994; Christina & Bjork, 1991; Schmidt & Bjork, 1992). In concordance with this view, the definition of learning that is used here is "the development of some relatively permanent capability for responding" (Schmidt & Bjork, 1992, p. 208). Experiments on the spacing effect typically end with a final retention test that is given after a sufficiently long interval to ensure that any temporary effects from the experimental manipulations have dissipated.

Although the spacing effect refers to the advantage in learning of spaced practice over massed practice, the effect is often reversed during acquisition and immediately following acquisition (e.g., Bregman, 1967; Lee & Genovese, 1988; Lee & Magill, 1983; Peterson et al., 1962; Peterson et al., 1963; Shea, Kohl, & Indermill, 1990; Shea & Morgan, 1979). The difference between the spacing effect obtained immediately following acquisition and the spacing effect obtained after a delay has been termed the "proportionality rule" by Glenberg and Lehmann (1980). According to Glenberg and Lehmann, "when the retention interval is short relative to the spacing of the repetitions, performance is negatively correlated with repetition spacings; when the retention interval is long relative to the spacing intervals, performance is positively correlated with the spacing of repetitions." (p. 528)

The proportionality rule was first demonstrated by Peterson et al.
In their study, Peterson et al. provided subjects with two presentations of words paired with numbers. The specific word-number presentations were separated by either 0 sec (Massed condition) or 8 sec (Spaced condition). Following the second presentation, the subjects were prompted for the correct numeric response to each of the words. The retention interval was either 2, 4, 8, or 16 sec. Although the massed presentations led to better recall at the 2- and 4-sec retention intervals, the spaced presentations led to better recall at the 8- and 16-sec intervals.

The study by Peterson et al. (1962) exemplifies that the spacing effect is dependent upon the interaction between the acquisition and retention intervals. As such, the spacing effect is among a variety of manipulations in which conditions that impair performance during training or acquisition, are optimal for long-term retention and performance.

The finding that learning improves the greater the temporal distribution of the repetitions is often called the Melton lag effect (or simply the "Melton" or "lag" effect) due to a seminal study by Melton (1967). In his study, Melton presented subjects with a series of words, some of which occurred twice in the list with either 0, 2, 4, 8, 20, or 40 intervening words between repetitions. Melton found performance improved as a function of the number of intervening words separating the presentations of the repeated words. However, as discussed above with regards to the proportionality rule, the relationship between the spacing of presentations and retention is not always monotonic (see Glenberg, 1976; Hintzman, 1974; Melton, 1970; Peterson, Hillner, & Saltzman, 1962; Peterson, Wampler, Kirkpatrick, & Saltzman, 1963). A systematic investigation of
the conditions leading to a monotonic lag effect has yet to be done (Greene, 1992). The distinction between the lag effect and the spacing effect is that the former compares the spacings of one or more intervening items while the latter compares massed practice (spacings of zero) with distributed practice (all spacings greater than zero).

During the past century, massed versus spaced presentations has become one of the most researched independent variables in the psychological literature, with well over 300 investigations (Bruce & Bahrick, 1992). The spacing effect attracted considerable attention during the 1960's and 1970's (for reviews see Crowder, 1976; Hintzman, 1974; Melton, 1970; Glenberg, 1979) and interest has continued to the present (for more recent reviews, see Dempster, 1988; Glenberg, 1992; and Lee & Genovese, 1988). The effect has been described as one of the most replicable and remarkably robust phenomena in experimental psychology (Dempster, 1987a. 1988). The robustness of the spacing effect is shown by the fact that, in some instances, spaced practice has led to over twice the advantage in learning of massed practice (e.g., Bahrick & Phelps, 1987). On the other hand, two massed practices have not been appreciably better than a single practice (e.g., Glenberg, 1979; Glenberg & Lehmann, 1980; Madigan, 1969; Melton, 1970; Peterson et al., 1963; Rothkopf & Coke, 1966). Furthermore, the spacing effect is ubiquitous in scope: the phenomenon has been observed in essentially every traditional learning paradigm, with a wide array of materials and methods spanning both the verbal learning and motor skill domains (Dempster, 1987a, 1988; Dempster & Farris, 1990).
Generality of the Spacing Effect

Although my goal is not to provide a detailed review of the 300+ studies reported in the literature documenting the spacing effect, in this section I provide an overview of some of the different methods and domains with which the spacing effect has been investigated to demonstrate the generality of the effect.

Verbal Learning and Memory

A review of the verbal learning literature shows that the spacing effect has been found using virtually every memory task (for reviews, see Crowder, 1976; Glenberg, 1992; Hintzman, 1974). Memory tasks in which the spacing effect has been found include free recall, in which the subject's task is to retrieve the information in any order without being provided with cues or hints; recognition testing, in which the subject is provided with the original studied information along with some new unstudied information and the subject's task is to recognize the original information; and frequency estimations, in which the subject's task is to provide an estimation of the number of times the different items had been presented during the acquisition phase. The spacing effect has also been found using continuous paired-associate learning, in which pairs of items are presented continuously, with new pairs interspersed with additional presentations or tests on old pairs. The items are tested in a cued recall fashion in which one of the items from the pair is given and the subject's task is to retrieve the item that had been paired with that item previously.

The spacing effect has been found using implicit memory tasks as well. Implicit memory tasks are indirect tests of memory in which retention for
a specific past experience is measured using a new task during which the subject does not consciously retrieve the original experience. Implicit memory tasks in which the spacing effect has been found include homophonic word spelling, word-fragment completion, and perceptual identification (Greene, 1990). In homophonic word spelling, pairs of words are presented auditorily, with the second word in the pair corresponding to the less common meaning of a homophone (e.g., taxi-fare). Subjects are then given a spelling test on single words, including the homophones from the previous phase of the experiment. The number of times the less common spelling is provided for those words that had been presented earlier is compared to a baseline or control condition for those words not presented earlier. Word fragment completion tests can be used as either explicit or implicit memory tasks. In implicit memory tasks, the word fragments are used to obtain a measure of priming from the previous exposure to the information, whereas in explicit memory tasks the task becomes one of cued recall in which the fragments are provided as cues for the retrieval of the studied information. Both implicit and explicit word fragment tasks have been found to show spacing effects. Finally, perceptual identification tasks begin with the visual presentation of a list of words, followed by a perceptual task in which words are flashed quickly on a screen and the subject's task is to identify each of the words. The number of words identified is compared between those words that had been presented previously and those words that had not.

Furthermore, the spacing effect has been found for both visually and auditorily presented information (Melton, 1970). The materials used to
demonstrate the effect have also varied widely, including unrelated words (for a review see Greene, 1989b), foreign language vocabulary (e.g., Bahrick, 1979; Bloom & Shuell, 1981; Feuge, 1976; Rea & Modigliani, 1988), vocabulary definitions (e.g., Dempster, 1987b), textbook chapters (Reder & Anderson, 1982), and lectures (Smith & Rothkopf, 1984).

Although the majority of the literature has operationalized distributed practice by the length of interval between successive trials (e.g., the lag between repetitions), researchers have also demonstrated the effect by manipulating the amount of practice per day across numerous learning sessions (e.g., Smith & Rothkopf, 1984). Furthermore, spacing effects have been found across a wide range of subject populations (for a review, see Glenberg, 1992). Additionally, the effect has been found to be remarkably durable, as demonstrated by Bahrick & Phelps (1987), in which a retention test was given 8 yrs after original training.

Motor Skill Acquisition

The effects of spacing in the verbal learning literature are analogous to blocked versus random practice effects in the motor skill literature (for reviews, see Annett, 1979; Lee & Genovese, 1988; Magill & Hall, 1990). Specifically, a blocked practice schedule is analogous to massed practice in that all trials for a given task are completed prior to moving on to practice on the next task. A random practice schedule, on the other hand, is analogous to spaced practice in that the order of practice on the various tasks is randomized such that no particular task is practiced on successive trials.

Among the variety of tasks in which distributing practice has been
investigated, spacing effects have been documented in the following motor skill domains: rapid, multiple-component arm movement drills (e.g., Lee & Magill, 1983; Shea & Morgan, 1979); anticipation timing skills, in which the subject's task is to press a response button at the same time that a target light goes on (e.g., Del Rey, 1982, 1989); maze tracing skills and computer-based tracking, in which a joystick is used to maneuver a cursor through a maze, (e.g., Jelsma & Van Merrienboer, 1989; Jelsma & Pieters, 1989a, 1989b); real-world sports skills (e.g., Goode & Magill, 1986); keyboard skills (Baddeley & Longman, 1978); and military tasks (for a review, see Hagman & Rose, 1983).

Many researchers note that the fact that parallel findings have been demonstrated across both the verbal learning domain and the motor skill domain indicates that similar phenomena are being witnessed and that general principles for training can be deduced (Goode & Magill, 1986; Lee & Magill, 1983; Melton, 1970; Schmidt & Bjork, 1992). Due to the ubiquity of the effect, the definition of the spacing effect is fairly general (i.e., the material, method, procedure, etc. is, for the most part, left unspecified).

"Test-Spacing" Effect

Often, in the real world, rather than determining the optimum distribution of study presentations, the question is how to optimize practice retrieving the information. Although most of the research on the spacing effect has dealt with the spacing of presentations, a small number of studies have dealt with the spacing of tests (e.g., Landauer & Bjork, 1978; Rea & Modigliani, 1985; Siegel & Misselt, 1984). It has repeatedly been shown
that tests improve memory (e.g., Izawa, 1971; Whitten & Bjork, 1977). In fact, successfully retrieving information from memory is more potent as a learning event than simply presenting the information again (e.g., Hagman & Rose, 1983; Hogan & Kintsch, 1971; Landauer & Bjork 1978). Landauer and Bjork (1978) determined the optimum distribution of retrieval practice to consist of an expanding schedule such that the first test is given soon after the original presentation, followed by tests at successively longer delays. As Bjork (1988) stated, "Each successive retrieval helps insure a successful retrieval after the next (longer) interval, and as the interval gets longer each retrieval becomes more potent as a learning event." (p. 399) In a study by Rea and Modigliani (1985), children's retention of multiplication facts under an expanded retrieval practice schedule was about twice that found under conditions of massed practice.

There is some dispute among researchers over which phenomena should be included under the rubric of the "spacing effect." Dempster (1988) notes that "the spacing effect is one of a family of similar, though less thoroughly investigated, phenomena that are often confused in the literature"--amongst which he includes the so-called "test-spacing effect" (p. 630). However, for the purposes of this dissertation, a general definition of the spacing effect is adopted; one that includes spacing effects for both presentations and tests. One reason for this decision is that the theory tested in the current research applies regardless of whether the spacing is of presentations or tests--although different magnitudes of effect are predicted for the former and the latter. In fact, test-spacing effects
have been explained in terms of successful retrieval practice acting as repetitions (Landauer, 1969; Landauer & Bjork, 1978; Whitten & Bjork, 1977). Moreover, it has been argued that the principal mechanism underlying the standard spacing effect might be differences in the nature of the retrieval processes that take place at the time of massed versus spaced repetitions (R. A. Bjork, personal communication, June 27, 1996; Cain & Wiley, 1939; Schmidt & Bjork, 1992). For example, Schmidt and Bjork (1992) state that "the spacing of repetitions may prevent superficial massed rehearsal." (p. 212) For these reasons, the term "spacing effect" will be used here to refer to the increase in retention for items practiced (through additional test and/or study episodes) at nonadjacent times.

Current Theories

The ubiquity of the spacing effect makes it difficult to find an explanation that holds across all of the conditions in which the effect is found. Difficulty in finding an adequate explanation also stems from the fact that the effect is considered to be an "anomaly" or counterintuitive for the following reasons: (1) The spacing effect violates Jost's (1897, as described in Hintzman, 1974) law of recency since the first presentation is typically more recent (in terms of the retention test) under massed conditions than under spaced conditions (see Figure 1); (2) It violates the total-time law (Bugelski, 1962; Cooper & Pantle, 1967) that retention depends on the overall time of study and not on how that time is distributed; and (3) As is often the case, massed repetitions are fairly infrequent (compared to spaced repetitions) and should therefore lead to a von Restorff or isolation effect, increasing rather than decreasing the
Massed Condition:

Presentation 1, Presentation 2...................Test

Spaced Condition:

Presentation 1............................Presentation 2...................Test

Figure 1. Temporal sequence of presentations and test under typical massed and spaced conditions. The time sequence goes from left to right and the dots (i.e., "...") represent time elapsed. Note that the second presentation and final retention test occur at the same time in both the massed and spaced conditions.
retention of the massed items. (For further discussions on the
counterintuitiveness of the spacing effect, see Bjork, 1979; Crowder, 1976;
Dempster, 1988; Hintzman, 1974.)

The myriad of explanations proposed to account for the spacing effect
can best be classified into two general types of theory: deficient processing
theories and encoding variability theories.

**Deficient Processing Theories**

According to deficient processing theories, individuals do not fully
process the repetitions of an item when it is presented under massed
conditions compared to when it is presented under spaced conditions. This
statement is supported by findings indicating that few details about the
second presentation of an item are remembered by the subject when the
item is presented under massed conditions (Hintzman, Block, & Summers,
1973). Other evidence in favor of deficient processing theories comes
from studies involving the use of overt rehearsals, in which it has been
found that more overt rehearsals are given to the second presentation of an
item under spaced than under massed conditions (e.g., Ciccone &
Brelsford, 1974). Further, in self-pacing studies, when given the chance to
pace themselves through a series of items presented on slides, subjects
spend less exposure time overall on massed items than they spend on spaced
items (e.g., Shaughnessy, Zimmerman, & Underwood, 1972). Additional
support for deficient processing theories comes from studies of pupil
dilation, in which it has been found that subjects' eyes dilate more (which is
presumed to be an indication of increased cognitive processing) when
spaced items are presented than when massed items are presented (e.g.,
Magliero, 1983). Moreover, dual-task response times seem to support the deficient processing notion, in that it has been shown that subjects respond slower to an auditory tone sounded during the presentation of spaced items than they do during the presentation of massed items (e.g., Johnston & Uhl, 1976). The underlying premise here is that subjects will be slower to respond when they are processing an item.

One problem, however, for deficient processing theories is that the type of memory test used appears to be important in explaining the spacing effect. For example, although spacing effects are not observed when a recognition test is given for incidentally-learned material (i.e., material that is presented without the subject's knowledge that the material will be tested later), spacing effects are observed when the same material is tested by free recall (see Greene, 1989a, 1990).

**Encoding Variability Theories**

According to encoding variability theories, the information remembered from each presentation of a repeated event differs from presentation to presentation. Assuming the information remembered differs more under spaced conditions than under massed conditions, the spacing effect can be explained by the increased probability of successfully retrieving an item from at least one of its presentations the greater the difference from each encoding. Examples of some of the various types of encoding variability explanations proposed to account for spacing effects include variations in the contextual information (e.g., Glenberg, 1979; Landauer, 1976; Melton, 1970) and variations in the interpretation or connotation of the items (e.g., Gartman & Johnson, 1972).
One problem for encoding variability theories is that they predict an increased probability of retrieving either of two once-presented items as the spacing between the two items increases; however this has not been found to be the case (e.g., Ross & Landauer, 1978).

Although numerous explanations, including deficient processing explanations and encoding variability explanations, have been proposed to account for the spacing effect, an adequate explanation—one that holds across the abundance of varied conditions—remains to be found. For reviews of different theories, see Crowder, 1976; Glenberg, 1992; Greene, 1989b, 1992; and Hintzman, 1974, 1976.

The New Theory of Disuse

Recently, Bjork and Bjork (1992) proposed a theory, referred to as the new theory of disuse, that accounts for a number of standard experimental findings, including the spacing effect. The assumptions of their theory can be summarized as follows (For a thorough discussion of the assumptions, see Bjork & Bjork, 1992, pp. 42-44):

1. Items in memory have two different "strengths," referred to as a storage strength and a retrieval strength. The storage strength of an item equates to how well learned the item is, but does not have a direct effect on the probability that the item will be retrievable from memory at a given time. The retrieval strength of an item equates to the item's current accessibility from memory and directly determines the probability that the item will be retrievable in response to a given cue.

An example of something that might have both a high storage strength and a high retrieval strength is a person's own name. On the other hand,
something that might have both a low storage strength and a low retrieval strength is a person's great grandmother's name. The storage strength and retrieval strength of a particular item, however, are not necessarily both high or both low. For example, something that might have a high storage strength and a low retrieval strength may be the zip code from a person's previous residence. The phone number at a person's new job might have a low storage strength and a high retrieval strength.

2. An item's storage strength increases as a function of the number of study or recall opportunities for that item. Furthermore, it is assumed that storage strength is never lost: it is characterized as a negatively accelerated monotonic function. That is, the increase in an item's storage strength from a study or test opportunity is believed to be a decreasing function of the item's current storage strength. The increase in an item's storage strength is also assumed to be a decreasing function of the item's current retrieval strength. In other words, "high retrieval strength is assumed to retard the accumulation of additional storage strength." (Bjork & Bjork, 1992, p. 43) (This last assumption plays a crucial role in the theory's ability to explain the spacing effect and will be discussed further in the next section.)

3. Unlike the capacity for items to be stored in memory, there is an overall limit on the number of accessible items from memory at any given time. As new items are learned, or the retrieval strengths of items in memory are increased, other items become less accessible.

4. Although an item's storage strength and retrieval strength increase as a function of both study and test opportunities on the item, it is the test
opportunities that is assumed to be the more potent learning event of the two. In other words, successfully retrieving an item from memory will lead to a greater increase in the item's storage strength and retrieval strength than will studying the item. In addition, for both study and test opportunities, the increase in the item's retrieval strength is believed to decrease the higher the item's current retrieval strength and increase the higher the item's current storage strength.

5. The decrease in an item's retrieval strength due to the learning or retrieval of other items in memory is believed to increase the higher the item's current retrieval strength and decrease the higher the item's current storage strength. In other words, "the gain and loss of retrieval strength are both negatively accelerated, and ... storage strength acts to enhance the gain and retard the loss." (Bjork & Bjork, 1992, p. 44) (This assumption is also important in explaining the spacing effect and will be returned to in the next section.)

Implications for the spacing effect. Bjork and Bjork's (1992) new theory of disuse can account for the spacing effect in that:

In general, spacing of repetitions results in higher storage strength than does massing of repetitions, which in turn slows the rate of loss of retrieval strength and, therefore, enhances long-term performance. Massing, however, can produce a higher initial level of retrieval strength, which, given a short enough retention interval, can result in a higher level of recall than that produced by spaced repetitions. (pp. 46-47)

Thus, massed practice leads to a more rapid accumulation of retrieval
strength (because not much is lost between trials) than does spaced practice. However, because increments in storage strength are a decreasing function of current retrieval strength, spaced practice yields higher increments in storage strength.

Based on the assumptions of the new theory of disuse (Bjork & Bjork, 1992), Bjork and Fritz (1994) derived mathematical equations (see Appendix) and generated a quantitative example to illustrate the dynamics of an item's storage strength and retrieval strength depending upon whether that item is practiced under massed or spaced conditions. The resulting graphs are shown in Figures 2 and 3. Figure 4 illustrates the retrieval strengths after each of the practice trials under massed and spaced training conditions, as well as the predicted retrieval strengths on a delayed retention test. The quantitative example by Bjork and Fritz parallels the results typically found in distribution of practice studies. Specifically, performance, as indexed by an item's retrieval strength, is better during the acquisition trials under massed conditions than it is under spaced conditions. However, on the final retention test, performance is better for an item acquired under spaced conditions than it is under massed conditions.

Moreover, Bjork and Fritz (1994) determined that the new theory of disuse makes an interesting prediction for relearning. Namely, the theory predicts that for an item that had been originally learned well, as indexed by a high storage strength, the spacing effect should not be found, or may even reverse itself. In other words, spaced practice should not lead to better retention than should massed practice in a relearning situation.
Figure 2. Retrieval strength (R) and storage strength (S) as a function of practice under massed original learning conditions. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
Figure 3. Retrieval strength (R) and storage strength (S) as a function of practice under spaced original learning conditions. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
Figure 4. Retrieval strength / performance for massed and spaced conditions during the acquisition trials and on the retention test during original learning. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
The logic behind this prediction is that if storage strength—which, according to the theory, is never lost—is already high from previous learning, then massed relearning can have performance advantages over spaced relearning for a considerable period of time. That is, high storage strength will help to retard the loss of retrieval strength, therefore enhancing long-term performance.

Figures 5 and 6 replicate the mathematical example generated by Bjork and Fritz (1994) to show an item's respective storage strengths and retrieval strengths during relearning under massed and spaced trials. In their example, the item is originally well learned, as indexed by a high storage strength at the beginning of the relearning trials. Figure 7 illustrates the retrieval strengths after each of the relearning trials under massed and spaced conditions, as well as the predicted retrieval strengths on a retention test. As can be seen in the figure, there is no longer an advantage on the final retention test for an item practiced under spaced conditions. In fact, performance is slightly better on the delayed test for an item practiced under massed conditions.

Rationale for the Present Research

The present research explores the effects of distributing practice during relearning in order to test Bjork and Fritz's (1994) prediction, based on the new theory of disuse (Bjork & Bjork, 1992), that spacing practice is not important for relearning. As Hintzman (1974) states, "The spacing effect is sufficiently general that failures to obtain it under the appropriate conditions are, potentially, of considerable theoretical importance." (p. 79) From an applied standpoint, it is important to determine whether principles
Figure 5. Retrieval strength (R) and storage strength (S) as a function of practice under massed relearning conditions. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
Figure 6. Retrieval strength (R) and storage strength (S) as a function of practice under spaced relearning conditions. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
Figure 7. Retrieval strength / performance for massed and spaced conditions during the acquisition trials and on the retention test during relearning. From "Reinforcing and Competitive Dynamics in the Maintenance of Knowledge," by R. A. Bjork and C. O. Fritz, 1994, August, in H. Bahrick (Chair), The maintenance of knowledge, symposium conducted at the Practical Aspects of Memory Conference, College Park, Maryland. Adapted with permission of the author.
that hold during original learning will generalize to relearning. As Magill and Hall (1990) note in their review of blocked versus random practice in motor skill acquisition, whether the advantage of random practice over blocked practice is a learning phenomenon tied only to the early stages of learning is not known and merits investigation.

**Considerations for Selecting the Task and Design of the Present Research**

There are two main requirements for a valid test of Bjork and Fritz's (1994) prediction on the effects of distributing practice during relearning. First, it is important for the item's storage strength to be at a fairly high level going into relearning. Although it is impossible to obtain a direct measure of storage strength, the item's accumulated storage strength should be near asymptote. Second, it is important to avoid a quick asymptote in the retrieval strength functions during relearning. In other words, no differences may be found between the massed and spaced conditions during relearning--not because the conditions do not differ, but rather because performance during relearning may asymptote quickly.

Annett (1979) cites evidence supporting the commonly-held notion that motor skills are better retained than verbal skills and states that, "Well-learned motor skills are generally well retained in the sense that they are rapidly relearned in only a few trials even after extensive periods without practice." (p. 238) Moreover, Melton (1970) notes that research interest on the spacing effect turned away from perceptual-motor skill tasks towards verbal learning because "it was recognized that the basic theoretical issues could be examined in a more analytic fashion in simple associative learning tasks involving units, in what has been quite
superficially described as 'rote learning'" (pp. 596-597). In line with these points, the decision was made to use a memory task rather than a motor task.

Additional considerations for selecting the task and design of the present research were a direct result of some pilot research. The pilot research helped provide insight on the possible pitfalls associated with designing an appropriate study to explore the effects of distributing practice during relearning. In the next section, the design of the pilot experiment is described, as well as a discussion of one of the problems encountered that helped lead to the present series of experiments.

**Pilot research.** The study examined the learning of foreign language vocabulary under spaced and massed conditions during original learning and relearning. Each subject learned Hebrew (transliterated) words with their English equivalents under both massed and spaced conditions. Subjects were given a retention test at the end of the first learning session and then returned 2 weeks later for relearning of the Hebrew-English word pairs. The relearning of the word pairs was given under both massed and spaced conditions. Subjects were given a final retention test 1 week after the relearning session.

The traditional spacing effect, in which spaced practice yields better (long-term) performance than massed practice, was expected to be found during original learning. Of primary interest was whether Bjork and Fritz's (1994) prediction, based on the assumptions of the new theory of disuse (Bjork & Bjork, 1992), would be confirmed--such that the spacing effect would disappear or reverse itself on the final retention test.
The results of the experiment after the original learning session were not surprising in that more of the English words were recalled in response to each of their appropriate Hebrew equivalents for the items that had been learned under spaced conditions than for the items that had been learned under massed conditions. However, there was a problem in the data because it appeared that learning was still taking place during the relearning session. This was indicated by the fact that performance after original learning (as measured by the first trial during the relearning session) was substantially lower than performance after the relearning session (as measured by the final retention test).

One criticism of comparing performance during the first trial of the relearning session to that of the final retention test is that the retention interval for the former measure was 2 weeks, but the retention interval for the latter measure was only 1 week. However, the forgetting of verbal material is believed to asymptote fairly quickly. For instance, Krueger (1929) found retention after learning to decrease dramatically after the first day or so, and then to remain relatively unchanged after the first week (see also Woodworth & Schlosberg, 1954). Converging evidence on the predominance of rapid forgetting of verbal material to occur within the first 24 hrs is cited by Slamecka and McElree (1983). Therefore, the possibility that additional learning occurred during the relearning session cannot be ruled out. This posed a damaging blow to the study because the new theory of disuse (Bjork & Bjork, 1992) makes opposing predictions for original learning and relearning. As a result, the study failed to provide an adequate test of the spacing effect during relearning.
The main failure of the pilot study was its mission to carry-out both original learning and relearning. Namely, the study made clear the difficulty, within the constraints of most laboratory settings, to reach an adequate level of original learning, such that additional learning will not occur during the relearning phase of the research.

Present research. In the present research, a different approach to studying the distribution of practice during relearning was taken. In the first part of Experiment 1 (referred to as Experiment 1a), the issue of the spacing effect during relearning was explored using the task of mental arithmetic. It was assumed that all college students should have learned mental arithmetic well enough in grade school (i.e., storage strength should be close to asymptote). However, based on anecdotal evidence (and, in fact, borne out in the present studies), the ease with which a person can solve mental arithmetic problems is a skill that substantially slows with periods of disuse (e.g., as measured by reaction times). As a result, the task of mental arithmetic lends itself nicely to an experimental design requiring only one experimental session. That is, the ease with which subjects solve mental arithmetic problems can be explored in a relearning session consisting of massed and spaced practice conditions. Another benefit of using the task of mental arithmetic is that the dependent variable (i.e., reaction times) should serve to avoid potential problems with performance asymptoting quickly during relearning.

In the second part of Experiment 1 (referred to as Experiment 1b), a different group of subjects, serving as a comparison group to Experiment 1a, were given an analogous task—mental arithmetic involving letters
rather than numbers—in order to obtain a measure of the spacing effect under massed and spaced practice conditions during original learning. It should be noted that substantial spacing effects have been found in the retention of multiplication facts taught to grade school children (Rea and Modigliani, 1985). Experiment 1b provides a test of the spacing effect on an analogous task.

Experiment 2 was similar to Experiment 1a, except for the instructions given to the subjects during the experiment. The main reason for attempting to replicate the first part of Experiment 1 was because there was the hint of a difference in the accuracy data between the massed condition and one of the spaced conditions. The instructions for Experiment 2 placed more stress on accuracy in order to avoid problems associated with a speed/accuracy trade-off.

Experiment 1a

In Experiment 1a, the effects of spacing during relearning was investigated using a task in which original learning presumably occurred years before the experimental relearning session. The ease with which subjects were able to solve single-digit multiplication problems (as measured by reaction times) was determined under massed and spaced relearning conditions.

In the experiment, subjects were given simple multiplication problems to solve on a computer. Each of the critical problems was presented four times under massed or spaced conditions. The presentation of a problem consisted of a test-study trial, such that the problem was first presented without its solution and the subject was cued to provide the appropriate
answer. The problem was then shown with its correct solution. The main attraction of using test trials as opposed to simply presenting both the problem and solution during each of the presentations stemmed from the following considerations: Mental arithmetic is typically learned in the classroom using drill procedures (e.g., timed tests) as opposed to simply studying the material. A test trial is only effective to the degree that the subject is able to successfully retrieve the information. If the subject is unable to successfully retrieve the information on the first test trial, it is unlikely that he/she will retrieve the information on any of the remaining trials. As a result, it was decided to present the appropriate response following each of the test trials. Although it was assumed that accuracy should be near 100% on the task used in this experiment, the study part of the test-study trials were included to parallel the procedure used in Experiment 1b (in which a more difficult task was employed).

Specifically, subjects received practice solving simple numeric multiplication problems under three different conditions: (1) A massed condition in which all of the practice on a particular problem occurred consecutively (i.e., lag = 0), (2) a spaced condition in which the practice on a particular problem was spaced in an expanded fashion, such that first there were no intervening problems between practices, then there were 4 problems, followed by 8 (i.e., lag = 0, 4, 8), and (3) an intermediate spacing between the massed condition and the expanded-spacing condition in which there was a uniform spacing of one intervening problem between each practice on the particular problem (i.e., lag = 1).

Following the relearning phase, the subjects were given a retention test
on the multiplication problems. According to Bjork and Fritz (1994) and the new theory of disuse (Bjork & Bjork, 1992), a traditional spacing effect should not be found for the numeric multiplication problems. The rationale behind such a prediction is that mental arithmetic (including simple multiplication problem solving) is a well-learned skill with a presumably high level of storage strength coming into the experiment, that will help retard the loss of retrieval strength during relearning under both massed and spaced conditions. Thus, a similar pattern of results is expected to that shown from the mathematical simulation of relearning in Figure 7.

Method

Subjects

Twenty-four Rice University undergraduates participated to satisfy a course requirement.

Design

The experiment comprised a one-way within-subjects design. The critical manipulation consisted of the interpresentation lag during the relearning phase. Lag is defined as the number of other multiplication problems intervening between repetitions. Multiplication problems were presented at three different lags: Massed (lag = 0), Spaced (lag = 1), and Expanded (lag = 0, 4, 8).

Materials

Simple multiplication problems were selected if they met the following two criteria: (1) The numbers being multiplied must be single-digit numbers; and (2) The solution to the problem must be a double-digit number (e.g., $2 \times 5 = 10$). (As a result of the second of these two criteria,
the numbers '0' and '1' were not included in the problems, but could exist in the solutions to the problems.) These criteria resulted in a total of 32 multiplication problems. The 18 most difficult problems (based on the experimenter's discretion) were designated critical, and the remaining 14 problems noncritical. The noncritical, or filler, problems were used both to fulfill the requirements of the lag sequences and to serve as primacy and recency buffers (see Tables 1 and 2).

**Study lists.** Each of the critical multiplication problems received four presentations, with the filler problems also presented from one to four times. The critical problems and noncritical buffer problems were indistinguishable from the subject's perspective.

Each presentation of a multiplication problem consisted of a test-study trial, such that first the multiplication problem test cue was given on a computer followed by the problem-solution pairing for study. The test cue was presented for 5 sec, during which time the subject was to type the solution to the problem (e.g., $3 \times 7 = ?$). The length of the test trials was uniform, such that even after a response was typed, the problem remained on the screen for the full time. The problem-solution pairing was then shown for 3 sec during the study portion of the test-study trial.

As can be seen in Table 3, three different study lists were required to complete the counterbalancing. Each of the lists contained the same 18 critical multiplication problems, but the lag at which each was repeated varied across lists. For the purpose of counterbalancing the study lists, the 18 numeric multiplication problems were divided into three sets of 6 (Sets A, B, and C). For a given list, one of the sets of problems was presented at
Table 1

Critical Multiplication Problems for Experiments 1a and 2

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Set A</th>
<th>Set B</th>
<th>Set C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3 x 7 = 21</td>
<td>5 x 8 = 40</td>
<td>7 x 8 = 56</td>
</tr>
<tr>
<td>2</td>
<td>8 x 9 = 72</td>
<td>7 x 9 = 63</td>
<td>9 x 9 = 81</td>
</tr>
<tr>
<td>3</td>
<td>5 x 7 = 35</td>
<td>6 x 7 = 42</td>
<td>4 x 7 = 28</td>
</tr>
<tr>
<td>4</td>
<td>4 x 9 = 36</td>
<td>4 x 8 = 32</td>
<td>5 x 9 = 45</td>
</tr>
<tr>
<td>5</td>
<td>6 x 9 = 54</td>
<td>5 x 6 = 30</td>
<td>6 x 8 = 48</td>
</tr>
<tr>
<td>6</td>
<td>4 x 6 = 24</td>
<td>3 x 9 = 27</td>
<td>3 x 6 = 18</td>
</tr>
</tbody>
</table>
Table 2

*Filler Multiplication Problems for Experiments 1a and 2*

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Problem and Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 x 9 = 18</td>
</tr>
<tr>
<td>2</td>
<td>4 x 5 = 20</td>
</tr>
<tr>
<td>3</td>
<td>3 x 4 = 12</td>
</tr>
<tr>
<td>4</td>
<td>6 x 6 = 36</td>
</tr>
<tr>
<td>5</td>
<td>8 x 8 = 64</td>
</tr>
<tr>
<td>6</td>
<td>2 x 8 = 16</td>
</tr>
<tr>
<td>7</td>
<td>2 x 6 = 12</td>
</tr>
<tr>
<td>8</td>
<td>2 x 5 = 10</td>
</tr>
<tr>
<td>9</td>
<td>3 x 5 = 15</td>
</tr>
<tr>
<td>10</td>
<td>7 x 7 = 49</td>
</tr>
<tr>
<td>11</td>
<td>5 x 5 = 25</td>
</tr>
<tr>
<td>12</td>
<td>2 x 7 = 14</td>
</tr>
<tr>
<td>13</td>
<td>3 x 8 = 24</td>
</tr>
<tr>
<td>14</td>
<td>4 x 4 = 16</td>
</tr>
</tbody>
</table>
Table 3
Counterbalancing of Multiplication Problems in Experiments 1a, 1b, and 2

<table>
<thead>
<tr>
<th>Study List</th>
<th>Massed: 0</th>
<th>Spaced: 1</th>
<th>Expanded Spaced: 0, 4, 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. A-C refers to sets of critical multiplication problems (see Tables 1 and 7).
each of the three different lag conditions. Across the three lists, each of the sets of problems occurred once in each of the lag conditions. Thus, each of the 18 critical numeric multiplication problems occurred equally often at each of the lags. This counterbalancing of the lag conditions was completed with three subjects (i.e., one subject for each of the three study lists). Of course, the division of the study lists into distinct sets was not apparent to the subjects.

The same sequencing was used for all of the study lists (see Table 4). The sequencing was developed with the constraint that the first presentation of a problem at a given lag not be able to hold a serial position immediately adjacent to the first presentation of another problem at the same lag. Further, the median serial position (i.e., the average serial position of the second and third presentations) amongst the Massed, Spaced, and Expanded critical problems was controlled (median serial position of 49 for all three of the experimental conditions; see Table 5).

**Retention test.** The retention test was presented on a computer in a format analogous to the test trials in the relearning phase. However, no feedback or study followed any of the test trials. Each of the 18 critical multiplication problems was randomly selected by a Hypercard computer program and presented for 5 sec, and the subject was cued to type the solution to the problem. Thus, the order in which the 18 critical problems were tested was determined randomly, and was different for each of the subjects.

**Distractor task.** Form IV of the Wonderlic Personnel Test (1992) was administered during part of a 30-min retention interval between the
Table 4

Sequencing of Massed (M), Spaced (S), Expanded (E), and Filler (F) Trials for Study Lists in Experiments 1a, 1b, and 2

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F1</td>
<td>27 M2</td>
<td>53 F8</td>
<td>79 S5</td>
</tr>
<tr>
<td>2</td>
<td>F1</td>
<td>28 M2</td>
<td>54 S4</td>
<td>80 E6</td>
</tr>
<tr>
<td>3</td>
<td>F1</td>
<td>29 M2</td>
<td>55 F8</td>
<td>81 S5</td>
</tr>
<tr>
<td>4</td>
<td>F2</td>
<td>30 M2</td>
<td>56 M4</td>
<td>82 F11</td>
</tr>
<tr>
<td>5</td>
<td>F2</td>
<td>31 E2</td>
<td>57 M4</td>
<td>83 S5</td>
</tr>
<tr>
<td>6</td>
<td>F2</td>
<td>32 S3</td>
<td>58 M4</td>
<td>84 M6</td>
</tr>
<tr>
<td>7</td>
<td>F2</td>
<td>33 F4</td>
<td>59 M4</td>
<td>85 M6</td>
</tr>
<tr>
<td>8</td>
<td>E1</td>
<td>34 S3</td>
<td>60 E3</td>
<td>86 M6</td>
</tr>
<tr>
<td>9</td>
<td>E1</td>
<td>35 F5</td>
<td>61 E4</td>
<td>87 M6</td>
</tr>
<tr>
<td>10</td>
<td>M1</td>
<td>36 S3</td>
<td>62 E4</td>
<td>88 S6</td>
</tr>
<tr>
<td>11</td>
<td>M1</td>
<td>37 F6</td>
<td>63 E5</td>
<td>89 E6</td>
</tr>
<tr>
<td>12</td>
<td>M1</td>
<td>38 S3</td>
<td>64 E5</td>
<td>90 S6</td>
</tr>
<tr>
<td>13</td>
<td>M1</td>
<td>39 F6</td>
<td>65 F9</td>
<td>91 F11</td>
</tr>
<tr>
<td>14</td>
<td>E1</td>
<td>40 E2</td>
<td>66 F9</td>
<td>92 S6</td>
</tr>
<tr>
<td>15</td>
<td>S1</td>
<td>41 M3</td>
<td>67 E4</td>
<td>93 F12</td>
</tr>
<tr>
<td>16</td>
<td>F3</td>
<td>42 M3</td>
<td>68 F10</td>
<td>94 S6</td>
</tr>
<tr>
<td>17</td>
<td>S1</td>
<td>43 M3</td>
<td>69 E5</td>
<td>95 F13</td>
</tr>
<tr>
<td>18</td>
<td>S2</td>
<td>44 M3</td>
<td>70 M5</td>
<td>96 F13</td>
</tr>
<tr>
<td>19</td>
<td>S1</td>
<td>45 E3</td>
<td>71 M5</td>
<td>97 F13</td>
</tr>
<tr>
<td>20</td>
<td>S2</td>
<td>46 E3</td>
<td>72 M5</td>
<td>98 F13</td>
</tr>
<tr>
<td>21</td>
<td>S1</td>
<td>47 F7</td>
<td>73 M5</td>
<td>99 F14</td>
</tr>
<tr>
<td>22</td>
<td>S2</td>
<td>48 S4</td>
<td>74 E6</td>
<td>100 F14</td>
</tr>
<tr>
<td>23</td>
<td>E1</td>
<td>49 F7</td>
<td>75 E6</td>
<td>101 F14</td>
</tr>
<tr>
<td>24</td>
<td>S2</td>
<td>50 S4</td>
<td>76 E4</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>E2</td>
<td>51 E3</td>
<td>77 S5</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>E2</td>
<td>52 S4</td>
<td>78 E5</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Numbers following the trial-type refer to the problem number presented (see Tables 1, 2, 7, and 8).
Table 5

Median Serial Position in the Study Lists for the Experimental Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Serial Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd Presentation</td>
</tr>
<tr>
<td>Massed (Lag = 0)</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td>Average</td>
<td>49.00</td>
</tr>
<tr>
<td>Spaced (Lag = 1)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>90</td>
</tr>
<tr>
<td>Average</td>
<td>48.33</td>
</tr>
<tr>
<td>Expanded (Lag = 0, 4, 8)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>75</td>
</tr>
<tr>
<td>Average</td>
<td>47.00</td>
</tr>
</tbody>
</table>
relearning session and the retention test. The test was used to determine if any of the experimental effects correlate or interact with cognitive ability. **Procedure**

Subjects were tested in groups of three or fewer. The experimental session began with the relearning phase, during which the subjects were presented with numeric multiplication problems on a computer and they received practice solving them. Before the experiment proper, the subjects were given a short practice set of multiplication problems to ensure they were comfortable with the procedure. The practice set consisted of a couple of the noncritical problems.

During the test trials in the relearning phase, the subjects were prompted to enter the solution to each of the multiplication problems. The subjects were informed that the test trials were timed and that they should try to respond as accurately and quickly as possible. They were also informed that there was no penalty for guessing. Due to the fact that reaction time data was being collected, the subjects were not allowed to switch their answer once they typed it.

Following the relearning phase, the subjects were given distractor tasks to work on for 30 min. The first distractor task administered was the Wonderlic Personnel Test (1992), after which they were given a spatial ability test. The subjects were then given a retention test on the 18 critical multiplication problems. Once again the subjects were prompted to enter the solution to each of the problems and they were told that the same instructions applied as on the test trials from the relearning phase. The entire experiment lasted approximately 1 hr.
Results

The results are presented in three sections. The first section summarizes the subjects' performance during acquisition. In the second, and more important, section the data from the retention test is presented. The third section correlates subjects' performance on the Wonderlic Personnel Test (1992) with their performance on the retention test. The data in both of the first two sections are further summarized in two subsections, according to whether accuracy or reaction time is the dependent variable.

In all of the experiments, the Bonferroni method was used to control for familywise error. The Bonferroni correction entailed multiplying the p-values of the planned comparisons by three (because there were three experimental means and therefore three pairwise comparisons). Unless otherwise noted, none of the effects that were nonsignificant with the Bonferroni adjustment would have been significant without it. This is particularly relevant in Experiments 1a and 2, in which no difference between conditions was predicted.

Acquisition Test Trials

Accuracy. Figure 8 shows the average proportion of correct responses during acquisition, as well as during retention. As expected, accuracy on the numeric multiplication problems was quite high and was not affected by experimental condition.

Reaction times. Reaction times were defined as the time between the presentation of the problem and the subject's first keypress. The first keypress was used rather than the time for subjects to provide the complete
Figure 8. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1a.
response to control for random error due to typing speed and the distance on the keyboard between the two digits in the response.

Only the reaction time data from correct responses were used in the analyses. In most analyses, the median correct reaction time (in sec) for each subject in each condition is the dependent variable. Median reaction times were, in general, more stable than mean reaction times.

Figure 9 presents the mean of the median correct reaction times in each condition. As expected, subjects' times decreased over trials. An exception to this pattern is the increase in time in the Expanded condition from Trial 2 to Trial 3. The second trial is the fastest in this condition due to the spacing between the first and second trial having a lag of 0.

Improvement on performance during acquisition was measured by the difference between Trial 4 and Trial 1. As can be seen in Figure 9, performance improved most in the Massed condition, followed by the Spaced condition, and least in the Expanded condition. A 3 (Massed vs. Spaced vs. Expanded) x 2 (Trial 1 vs. Trial 4) repeated measures Analysis of Variance (ANOVA) on the median reaction times revealed a significant interaction, $F(2, 46) = 8.72, p = .001$. The main effect of experimental condition was also significant, $F(2, 46) = 6.69, p = .003$, as was the main effect of acquisition trial, $F(1, 23) = 116.73, p < .001$. Responses were reliably faster on the final acquisition trial than they were on the first acquisition trial in all three of the experimental conditions: Massed, Spaced, and Expanded, $t(23) = 7.58, 7.39,$ and $6.50$ respectively, $ps < .001$ (Bonferroni method).

An analysis on the final acquisition trial found a significant effect of
Figure 9. Average median reaction times (in sec) of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1a.
experimental condition, $F(2, 46) = 43.61$, $p < .001$. Subjects were faster on the final acquisition trial under Massed relearning conditions than they were under Spaced or Expanded relearning conditions, $t(23) = 6.10$ and $7.79$ respectively, $ps < .001$; and they were faster under Spaced relearning conditions than they were under Expanded relearning conditions, $t(23) = 4.79$, $p < .001$ (Bonferroni method).

**Retention Test**

**Accuracy.** As in the acquisition trials, accuracy on the numeric multiplication problems during retention was fairly high (see Figure 8). The ANOVA on the proportion of correct responses was not significant, $F(2, 46) = 2.89$, $p = .066$. Nor was there a significant difference on any of the planned comparisons: Massed vs. Spaced, $t(23) = 2.21$, $p = .111$, Massed vs. Expanded, $t(23) = 1.13$, $p = .810$, and Spaced vs. Expanded, $t(23) = 1.44$, $p = .492$ (Bonferroni method). However, despite the fact that there is not convincing evidence to claim a spacing effect on these data, there is a hint that performance in the Spaced condition may be better than performance in the Massed condition (as indicated by a significant $p$-value prior to the Bonferroni correction).

**Reaction times.** Boxplots of the mean and median correct reaction times during the retention test are presented in Figures 10a and 10b. The relative stability of the median reaction times compared to the mean reaction times, in addition to their similarity in pattern, can be seen in the boxplots. Neither Figure 10a nor Figure 10b show any evidence of a spacing effect. An analysis on the median correct reaction times found no hint of an effect of experimental condition, $F(2, 46) = 0.20$, $p = .816$. The
Figures 10a and 10b. Boxplots of the mean and median correct reaction times (in sec) for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 1a.
planned comparisons between conditions also found no significant differences: Massed vs. Spaced, \( t(23) = 0.16, p > .5 \), Massed vs. Expanded, \( t(23) = 0.46, p > .5 \), and Spaced vs. Expanded, \( t(23) = 0.66, p > .5 \) (Bonferroni method).

**Wonderlic Personnel Test**

Subjects' scores on the Wonderlic Personnel Test (1992) were calculated based on the number of correct responses (out of a possible 50). Table 6 shows the correlations of the subjects' Wonderlic scores with their performance on the retention test. In addition to correlating Wonderlic performance to median correct reaction times in each of the three experimental conditions, correlations were also obtained between Wonderlic performance and measures of spacing effects. Spacing effect measures were based on the difference in median correct reaction times between the Massed condition and each of the other two conditions (i.e., Spaced - Massed and Expanded - Massed).

As can be seen in Table 6, performance on the Wonderlic test did not correlate highly with performance on the retention test. Thus, there was not a significant relationship between cognitive ability, as measured by performance on the Wonderlic Personnel Test, and reaction time to respond (correctly) to the multiplication problems.

**Discussion**

The primary purpose of Experiment 1a was to test Bjork and Fritz's (1994) prediction, based on the new theory of disuse (Bjork & Bjork, 1992), that spacing practice is not important for relearning. Although it is impossible to ever "prove" the null hypothesis, the results were clearly in
Table 6

Intercorrelations Between Wonderlic Personnel Test Scores and Median Correct Reaction Times on the Retention Test in Experiment 1a

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Wonderlic</td>
<td>--</td>
<td>-.01</td>
<td>.11</td>
<td>.02</td>
<td>.15</td>
<td>.03</td>
</tr>
<tr>
<td>2. Massed</td>
<td>--</td>
<td></td>
<td>.71***</td>
<td>.75***</td>
<td>-.46*</td>
<td>-.76***</td>
</tr>
<tr>
<td>3. Spaced</td>
<td>--</td>
<td></td>
<td>.72***</td>
<td>.30</td>
<td>-.35</td>
<td></td>
</tr>
<tr>
<td>4. Expanded</td>
<td>--</td>
<td></td>
<td></td>
<td>-.11</td>
<td>-.14</td>
<td></td>
</tr>
<tr>
<td>5. Spaced - Massed</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td>.58**</td>
<td></td>
</tr>
<tr>
<td>6. Expanded - Massed</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
line with the predictions of the theory.

Given that simple numeric multiplication problems are presumed to be learned well in grade school, but that people may have a tendency to become "rusty" without much practice and therefore slowed-down in the speed with which they respond, reaction time to correctly respond to the problems was selected as the dependent variable of interest. The data support this notion of slowed reaction times due to being out-of-practice in that reaction times, for the most part, improved during the relearning phase of the experiment. In addition, subjects' reaction times were faster on the final retention test than they were at the beginning of the experiment (i.e., Trial 1 of the relearning session).

Accuracy in solving the problems, on the other hand, was predicted to be near perfect given the simplicity of the problems. However, there was a hint of a spacing effect, with performance in the Spaced condition exceeding that of the Massed condition. The possibility of a spacing effect in the accuracy data will be addressed in Experiment 2. But first, Experiment 1b was conducted to serve as a comparison to the present study, by examining the distribution of practice during the initial learning of mental arithmetic problems.

Experiment 1b

Experiment 1b served as a comparison for Experiment 1a. In the experiment, subjects practiced an analogous task to that used in the first experiment, namely they learned the solutions to a series of multiplication problems involving letters rather than numbers.

By randomly assigning letters from the alphabet to the numbers in each
of the multiplication problems from Experiment 1a, the alphabetic task serves as a formal equivalent to the numeric task, while remaining a task of rote learning. That is, there should be no transfer or interference from the numeric equivalent. Thus, the alphabetic task parallels rote learning of multiplication tables learned in grade school and provides a reasonable comparison for the effects of spacing during original learning.

According to the new theory of disuse (Bjork & Bjork, 1992), there should be an advantage for spaced practice on the alphabetically-represented multiplication problems. The rationale behind such a prediction is that the alphabetic multiplication problems should have relatively low (if any) storage strength at the beginning of the experiment, nor should much storage strength accumulate under massed learning conditions. Therefore, during retention, retrieval strength should be lost at an accelerated rate under massed conditions compared to spaced conditions. Thus, a similar pattern of results is expected to that shown from the mathematical simulation of original learning in Figure 4.

Method

Subjects

Twenty-four Rice University undergraduates participated to satisfy a course requirement. None had participated in Experiment 1a.

Design

As with Experiment 1a, the experiment comprised a one-way within-subjects design, with the critical manipulation consisting of the interpresentation lag during the learning phase. Alphabetically-symbolized multiplication problems were presented at three different lags: Massed (lag
= 0), Spaced (lag = 1), and Expanded (lag = 0, 4, 8).

Materials

Parallel to the materials used in Experiment 1a, six different letters from the alphabet (analogous to the digits 2-9 in the numeric problems in the previous experiment) were used to form the alphabetic problems. The solutions to the problems (i.e., the responses) remained in the form of a double-digit number (e.g., C x A = 12). Rather than incorporating a simple conversion from the numeric problems in the previous experiment to form the alphabetic problems for this experiment, the problems were constructed such that there was no systematic relationship between the problems and the solutions. This was important in order for the task to remain one of strictly rote learning. Thus, once again there was a total of 32 multiplication problems: 18 critical problems and 14 noncritical, or filler problems (see Tables 7 and 8).

Study lists. Each of the critical multiplication problems received four presentations, with the filler problems also presented from one to four times. The critical problems and noncritical buffer problems were indistinguishable from the subject's perspective.

Unlike Experiment 1a, the initial presentation of a multiplication problem provided both the problem and solution for 5 sec (e.g., C x A = 12). That is, the first presentation of each of the problems consisted only of a study trial, as opposed to being tested first. The subsequent presentations consisted of test-study trials. Each of the test-study trials consisted of the multiplication problem test cue followed by the problem-solution pairing for study. The test cue was presented for 8 sec, during
<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Set A</th>
<th>Set B</th>
<th>Set C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C x A = 12</td>
<td>C x S = 42</td>
<td>A x Y = 14</td>
</tr>
<tr>
<td>2</td>
<td>S x Y = 45</td>
<td>Y x T = 10</td>
<td>N x N = 30</td>
</tr>
<tr>
<td>3</td>
<td>C x T = 16</td>
<td>S x A = 48</td>
<td>C x F = 81</td>
</tr>
<tr>
<td>4</td>
<td>R x Y = 27</td>
<td>F x Y = 28</td>
<td>S x T = 17</td>
</tr>
<tr>
<td>5</td>
<td>N x T = 32</td>
<td>R x N = 72</td>
<td>R x A = 25</td>
</tr>
<tr>
<td>6</td>
<td>R x F = 21</td>
<td>F x A = 12</td>
<td>S x F = 56</td>
</tr>
</tbody>
</table>

Table 7
Critical Multiplication Problems for Experiment 1b
Table 8

**Filler Multiplication Problems for Experiment 1b**

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Problem and Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R x T = 16</td>
</tr>
<tr>
<td>2</td>
<td>S x N = 24</td>
</tr>
<tr>
<td>3</td>
<td>F x T = 30</td>
</tr>
<tr>
<td>4</td>
<td>F x F = 40</td>
</tr>
<tr>
<td>5</td>
<td>Y x Y = 15</td>
</tr>
<tr>
<td>6</td>
<td>A x T = 49</td>
</tr>
<tr>
<td>7</td>
<td>N x A = 36</td>
</tr>
<tr>
<td>8</td>
<td>T x T = 18</td>
</tr>
<tr>
<td>9</td>
<td>C x N = 20</td>
</tr>
<tr>
<td>10</td>
<td>A x A = 52</td>
</tr>
<tr>
<td>11</td>
<td>N x Y = 36</td>
</tr>
<tr>
<td>12</td>
<td>N x F = 64</td>
</tr>
<tr>
<td>13</td>
<td>C x Y = 63</td>
</tr>
<tr>
<td>14</td>
<td>S x S = 35</td>
</tr>
</tbody>
</table>
which time the subject was to type the solution to the problem (e.g., C x A = ?). The length of the test trials was uniform, such that even after a response was typed, the problem remained on the computer screen for the full time. The problem-solution pairing was then shown for 5 sec during the study phase. (Note: The lengths of the test and study trials were longer in this experiment than in the previous experiment because of the increased difficulty of the task.)

As with Experiment 1a, three different study lists were required to complete the counterbalancing (refer to Table 3). Each of the lists contained the same 18 critical multiplication problems, but the lag at which each was repeated varied across lists. For the purpose of counterbalancing the study lists for the learning session, the 18 alphabetic multiplication problems were divided into three sets of 6 (Sets A, B, and C). For a given list, one of the sets of problems was presented at each of the three different lag conditions. Across the three lists, each of the sets of problems occurred once in each of the lag conditions. Thus, each of the 18 critical alphabetic multiplication problems occurred equally often at each of the lags. This counterbalancing of the lag conditions was completed with three subjects (i.e., one subject for each of the three study lists). Of course, the division of the study lists into distinct sets was not apparent to the subjects.

The same sequencing of trials as was used in Experiment 1a was used for this experiment (refer to Table 4).

Retention test. As with Experiment 1a, the retention test was presented on a computer in a format analogous to the test trials in the learning phase, with no feedback or study following any of the test trials. Each of the 18
critical multiplication problems was randomly selected by a Hypercard computer program and presented for 8 sec, and the subject was cued to type the solution to each problem. Thus, the order in which the 18 critical problems were tested was determined randomly, and was different for each of the subjects.

**Distractor task.** As with the previous experiment, Form IV of the Wonderlic Personnel Test (1992) was administered during part of the 30-min retention interval, in order to determine if any of the experimental effects correlate or interact with cognitive intelligence or ability.

**Procedure**

The procedure was similar to that used in Experiment 1a. Subjects were tested in groups of three or fewer. The experimental session began with the learning phase, during which the subjects were presented with alphabetic multiplication problems and solutions on a computer and they received practice solving them. The subjects received instructions explaining that they would be presented with multiplication problems represented alphabetically and that they were to learn the appropriate double-digit number for each of the problems. They were informed that the task was meant to be a simple rote learning task and that there was no systematic relationship between the letters in the problems and standard multiplication problems. Before the experiment proper, the subjects were given a short practice set of multiplication problems to ensure they were comfortable with the procedure. The practice set consisted of a couple of the noncritical problems.

During the test trials in the learning phase, the subjects were prompted
to enter the solution to each of the multiplication problems. The subjects were informed that the test trials were timed and that they should try to respond as accurately and quickly as possible. They were also informed that there was no penalty for guessing. Because reaction time data was being collected, the subjects were not allowed to switch their answer once they typed it.

Following the learning phase, the subjects were given distractor tasks to work on for 30 min. The first distractor task administered was the Wonderlic Personnel Test (1992), after which they were given a spatial ability test. The subjects were then given a retention test on the 18 critical alphabetic multiplication problems. Once again the subjects were prompted to enter the solution to each of the problems and they were told that the same instructions applied as on the test trials in the learning phase. The entire experiment lasted approximately 1 hr.

Results

The results are presented in three sections. The first section presents the acquisition data, the second section presents the retention data, and the final section presents correlations between subjects' performance on the Wonderlic Personnel Test (1992) and their performance on the retention test.

As in Experiment 1a, reaction times to respond to the problems were collected. However, due to the considerably lower accuracy rates in the present study, the analyses on the reaction time data are not very useful. The primary problem is that there are numerous cases in which subjects did not correctly respond to any of the problems in a particular
experimental condition, thereby providing no scorable reaction times for that condition. A secondary concern with the reaction time data is that subjects were, in a sense, penalized (i.e., have longer reaction times), for the problems for which they eventually responded correctly compared to those problems that they did not respond to in the allocated time. As a result, the reaction time analyses are not reported.

Acquisition Test Trials

Figure 11 presents the average proportion of correct responses for the three experimental conditions. There are no data for Trial 1 of acquisition because this trial consisted of only a study trial, rather than a test-study trial.

By the end of acquisition, there were substantial differences in the proportions of correct responses, $F(2, 46) = 49.30, p < .001$, with reliable differences between all of the experimental conditions. As can be seen in Figure 11, accuracy was higher on Trial 4 in the Massed condition than in either the Spaced or Expanded conditions, $t(23) = 2.70, p = .039$ and $t(23) = 8.00, p < .001$ respectively; while accuracy was higher in the Spaced condition than in the Expanded condition, $t(23) = 6.76, p < .001$ (Bonferroni method).

Retention Test

Subjects' performance was much lower on the final retention test than on the acquisition trials (see Figure 11). Boxplots of the proportion of correct responses during the retention test are presented in Figure 12. As can be seen in the boxplots, both the mean and median proportion correct of the Massed condition were substantially lower than those of the Spaced
Figure 11. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 1b.
Figure 12. Boxplots of the proportion of correct responses for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 1b.
and Expanded conditions. Also evident in the boxplots is the huge spread in the subjects' proportions of correct responses in the Expanded condition.

An analysis on the proportion of correct responses on the retention test revealed a significant effect of experimental condition, $F(2, 46) = 8.93, p = .001$. As expected, both the Spaced and Expanded learning conditions produced reliably better performance on retention than Massed learning, $t(23) = 3.01, p = .018$ and $t(23) = 3.82, p = .003$ respectively (Bonferroni method). The difference between the Spaced and Expanded conditions was not significant, $t(23) = 1.48, p = .459$ (Bonferroni method).

A measure of retention from the acquisition trials to the retention test was computed for each of the experimental conditions by comparing the proportion of correct responses on the final retention test to performance at the end of the learning phase (i.e., Trial 4 of acquisition). This comparison yielded retention measures of 22% for Massed learning, 40% for Spaced learning, and 80% for Expanded learning.

**Wonderlic Personnel Test**

Table 9 shows the correlations of subjects' scores on the Wonderlic Personnel Test (1992) with their performance on the retention test. Consistent with the findings of Experiment 1a, performance on the Wonderlic test did not correlate highly with performance on the retention test. There was no clear relationship between cognitive ability, as measured by subjects' scores on the Wonderlic Personnel Test, and the proportion of correct responses on the alphabetic multiplication problems.

**Discussion**

Experiment 1b succeeded in demonstrating a traditional spacing effect
Table 9

**Intercorrelations Between Wonderlic Personnel Test Scores and Proportion of Correct Responses on the Retention Test in Experiment 1b**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Wonderlic</td>
<td>--</td>
<td>-.01</td>
<td>.23</td>
<td>.10</td>
<td>.19</td>
<td>.10</td>
</tr>
<tr>
<td>2. Massed</td>
<td>--</td>
<td>.17</td>
<td>.25</td>
<td>-.62**</td>
<td>-.38</td>
<td></td>
</tr>
<tr>
<td>3. Spaced</td>
<td>--</td>
<td>.39</td>
<td>.67***</td>
<td>.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Expanded</td>
<td>--</td>
<td>.12</td>
<td>.80***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Spaced - Massed</td>
<td>--</td>
<td>.50*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Expanded - Massed</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
in an original learning situation analogous to the relearning situation of Experiment 1a. Specifically, both the Spaced and Expanded learning conditions lead to better retention performance of alphabetic multiplication problems than did the Massed learning condition. In fact, the proportion of correct responses on the retention test in the Expanded condition ($M = .45$) was more than twice that of the Massed condition ($M = .22$). This replicates the findings of Rea and Modigliani (1985) in which expanded practice lead to nearly twice the learning of multiplication facts taught to grade school children compared to massed practice.

Interestingly, the difference between performance on the final retention test of the Expanded ($M = .45$) and the Spaced ($M = .37$) conditions was not statistically significant. As can be seen in the boxplots in Figure 12, the proportion of correct responses for the Expanded condition was extremely variable, ranging from 0.0 to 1.0 (i.e., perfect performance). One reason for the lack of a reliable superiority of the Expanded condition may have to do with the interplay between the test and study portions of the test-study trials. Although Landauer and Bjork (1978) found an expanded sequence of trials to be best when the practice consisted of tests (i.e., retrieval practice), a uniform spacing sequence was slightly better for practice in the form of repetitions (i.e., presentations). Landauer and Bjork reasoned:

The longer the interval from initial presentation to test, the lower the probability of success, but the greater its benefit for long term retention....This idea suggests that the optimal schedule for test-type rehearsal would be a pattern of increasing intervals between successive tests. A first test-trial at a short interval would be likely to succeed and
strengthen an item sufficiently to survive a slightly longer interval that would yield a more effective second practice trial, etc. In contrast, when the information is repeated, very long intervals are not as much better than moderate intervals and very short intervals are worse (see e.g., Landauer, 1969) so uniform spacing should be better for repetition-type practice. (p. 626)

In the present studies, the study portion was added at the end of each of the trials for practical reasons in that test-study trials (often referred to as 'anticipation trials') are a common procedure for learning. For example, most rote learning drill procedures used to teach such topics as foreign language vocabulary or mathematical times tables typically involve feedback as to the correct response. Further, the study portions of the test-study trials were added due to the fact that if the retrieval of an item was unsuccessful then it would be unlikely that the remaining test-trials on that item would be successful without a reminder of the correct response.

Although test-study trials are good from a practical perspective, they complicate matters from a theoretical perspective. Not only do the effects from the test portions of the test-study trials have to be taken into account, but also the effects from the study portions, in addition to the interaction between the two (R. A. Bjork, personal communication, September 19, 1996). Basically, the test portions are likely to be more potent learning events for expanded practice than for spaced practice, while the study portions are likely to be less potent. Yet not much is known regarding the interplay between the two.

One study relevant to this question is that of Flores (1993). In her
experiment, subjects learned the German equivalents to a list of English words. The acquisition phase consisted of test-study trials. The critical part of the experiment comprised two factors: The first factor was whether the trials were spaced in a uniform or expanded fashion and the second factor was whether the average interval between trials was relatively short or long. These factors were combined orthogonally to yield four experimental conditions: (i) uniform-short interval \((\text{lag} = 3)\), (ii) expanded-short interval \((\text{lag} = 0, 1, 3, 8)\), (iii) uniform-long interval \((\text{lag} = 8)\), and (iv) expanded-long interval \((\text{lag} = 1, 3, 8, 20)\). When a relatively short interval was used, performance on the final retention test given 10 min following the final learning trial was better for the expanded sequence \((M = .72)\) than for the uniform sequence \((M = .67)\). However, when a relatively long interval was used, performance was slightly better for the uniform sequence \((M = .79)\) than for the expanded sequence \((M = .77)\).

In the present study, the average interval was relatively short (i.e., lag = '1' in the uniform-spaced condition and '0, 4, 8' in the expanded-spaced condition) compared to Flores (1993). Consistent with Flores' findings, an expanded sequence led to superior performance (as measured by the average proportion of correct responses), though not to a reliable extent, most likely due to the variability in the data.

Experiment 2

Experiment 2 was conducted because there was a hint of a spacing effect (between the Massed and Spaced conditions) in the accuracy data from Experiment 1a. The current experiment explored the possibility of a
speed-accuracy trade-off. The materials and procedure for Experiment 2 were identical to those of Experiment 1a, except that the importance of accuracy was further emphasized in the instructions.

**Method**

**Subjects**

Twenty-four Rice University undergraduates participated to satisfy a course requirement. None had participated in either of the first two experiments.

**Design**

As with Experiments 1a and 1b, the experiment comprised a one-way within-subjects design. The critical manipulation consisted of the interpresentation lag during the learning phase. Numeric multiplication problems were presented at three different lags: Massed (lag = 0), Spaced (lag = 1), and Expanded (lag = 0, 4, 8).

**Materials**

The materials were the same as those used in Experiment 1a.

**Procedure**

The procedure was identical to that of Experiment 1a, with the exception that the instructions were changed in order to convey the importance of accuracy to the subjects. Subjects were told that they should strive to be correct 100% of the time and that reaction time is only of secondary importance.

**Results**

As in the previous experiments, the results are summarized in three sections: an acquisition section, a retention test section, and a section
correlating retention performance with the Wonderlic Personnel Test (1992). The data in both of the first two sections are further divided into two subsections, according to whether accuracy or reaction time is the dependent variable. Although it is the second of these two dependent variables that is of interest in terms of replicating Experiment 1a, it is the first that is important in terms of determining whether the hint of a spacing effect in accuracy in Experiment 1a is "real" or dependable.

**Acquisition Test Trials**

**Accuracy.** As can be seen in Figure 13, accuracy was quite high and not affected by experimental condition.

**Reaction times.** Figure 14 presents the mean of the median correct times in each condition. The pattern of the acquisition trials is remarkably similar to that of Experiment 1a, with the subjects' times, for the most part, decreasing over trials.

As can be seen in Figure 14, improvement on performance (as measured by the difference between Trial 4 and Trial 1) was greatest in the Massed condition, followed by the Spaced condition, and least in the Expanded condition. A 3 (Massed vs. Spaced vs. Expanded) x 2 (Trial 1 vs. Trial 4) repeated measures ANOVA on the median reaction times revealed a significant interaction, $F(2, 46) = 13.88, p < .001$. The main effect of experimental condition was also significant, $F(2, 46) = 14.55, p < .001$, as was the main effect of acquisition trial, $F(1, 23) = 116.35, p < .001$. Responses were reliably faster on the final acquisition trial than they were on the first acquisition trial in all three of the experimental conditions: Massed, Spaced, and Expanded. $t(23) = 10.40, 8.54,$ and $6.34$. 
Figure 13. Average proportion of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 2.
Figure 14. Average median reaction times (in sec) of correct responses for Massed, Spaced, and Expanded conditions during the test trials of acquisition and on the final retention test in Experiment 2.
respectively, $p < .001$ (Bonferroni method).

An analysis on the final acquisition trial found that there were significant differences among the three conditions, $F(2, 46) = 44.85, p < .001$. Consistent with Experiment 1a, subjects were faster on the final acquisition trial under Massed relearning conditions than they were under Spaced or Expanded relearning conditions, $t(23) = 5.05$ and 7.93 respectively, $p < .001$; and they were faster under Spaced relearning conditions than they were under Expanded relearning conditions, $t(23) = 5.70, p < .001$ (Bonferroni method).

**Retention Test**

**Accuracy.** As can be seen in Figure 13, accuracy on the final retention test was quite high. The mean proportion of correct responses in all three conditions exceed those of Experiment 1a. Unlike the first experiment, the analysis on the proportion of correct responses revealed no hint of an effect of experimental condition, $F(2, 46) = 0.49, p = .614$. Nor was there a hint of an effect in any of the planned comparisons: Massed vs. Spaced, $t(23) = 1.00, p > .5$, Massed vs. Expanded, $t(23) = 0.32, p > .5$, and Spaced vs. Expanded, $t(23) = 0.69, p > .5$ (Bonferroni method).

**Reaction times.** Boxplots of the mean and median correct reaction times during the retention test are presented in Figures 15a and 15b. Neither Figure 15a nor Figure 15b show evidence of a spacing effect. An analysis on the median reaction times found no effect of experimental condition, $F(2, 46) = 0.71, p = .499$. As in Experiment 1a, there were no significant differences in any of the planned comparisons: Massed vs. Spaced, $t(23) = 0.93, p > .5$, Massed vs. Expanded, $t(23) = 1.16, p > .5$, and Spaced vs.
Figures 15a and 15b. Boxplots of the mean and median correct reaction times (in sec) for Massed, Spaced, and Expanded conditions during the final retention test in Experiment 2.
Expanded, $t(23) = 0.25$, $p > .5$ (Bonferroni method).

**Wonderlic Personnel Test**

Table 10 shows the correlations of subjects' scores on the Wonderlic Personnel Test (1992) with their performance on the retention test. As with Experiments 1a and 1b, performance on the Wonderlic test did not correlate highly with performance on the retention test. There was no clear relationship between cognitive ability and reaction times to respond to the multiplication problems.

**Discussion**

The findings of Experiment 2 essentially replicate those of Experiment 1a. When the data from the second experiment were scored in terms of the median reaction times to correctly respond to the problems on the retention test, no substantial differences were found amongst the three distribution of practice conditions. Thus, the findings support Bjork and Fritz's (1994) prediction, based on the new theory of disuse (Bjork & Bjork, 1992), that spacing practice is not important for relearning.

Furthermore, Experiment 2 succeeded in raising accuracy rates on the final retention test to an essentially equivalently high level in all three experimental conditions. Although it does not make much sense to statistically compare the accuracy rates due their being near ceiling, unlike Experiment 1a, there was no hint of an effect of experimental condition. The near-perfect accuracy rates in the present experiment reflect the expected level to which simple multiplication problems should be learned by college-level students.

Interestingly, in spite of the increased accuracy in this experiment
Table 10

Intercorrelations Between Wonderlic Personnel Test Scores and Median Correct Reaction Times on the Retention Test in Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Wonderlic</td>
<td>--</td>
<td>-.05</td>
<td>-.10</td>
<td>-.24</td>
<td>-.09</td>
</tr>
<tr>
<td>2.</td>
<td>Massed</td>
<td>--</td>
<td>.65**</td>
<td>.78***</td>
<td>.03</td>
<td>-.06</td>
</tr>
<tr>
<td>3.</td>
<td>Spaced</td>
<td>--</td>
<td>.79***</td>
<td>.78***</td>
<td>.41*</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Expanded</td>
<td>--</td>
<td>.40</td>
<td>.58**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Spaced - Massed</td>
<td>--</td>
<td>.59**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Expanded - Massed</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.  **p < .01.  ***p < .001.
compared to Experiment 1a, there was not a corresponding increase in the reaction times. Thus, there was not a tradeoff between accuracy and speed, as originally expected. The subjects in the present experiment performed accurately, and yet just as fast, if not faster, than the subjects in the first experiment.

General Discussion

The experiments reported here constitute an exploratory study of the effects of distributing practice during relearning. Specifically, they provide an initial test of the prediction by Bjork and Fritz (1994), based on the new theory of disuse (Bjork & Bjork, 1992), that spacing practice is not important for relearning.

Hintzman (1974) states, "The investigator who fails to obtain the [spacing] effect would probably be wise to suspect sampling error, ceiling effects, or a flaw in the experimental design." (p. 79) However, the current research is unusual in the sense that a spacing effect was predicted to not be found under conditions of relearning.

Theoretical Implications

The findings from this research are theoretically important in the sense that they provide merit to the idea that how well-learned the information or skill is (i.e., storage strength) needs to be taken into account. Specifically, the findings provide support for the dynamics of an interplay between the two types of strengths (i.e., storage strength and retrieval strength) specified by Bjork and Bjork (1992) in their new theory of disuse.

Based on the current research, the following conclusions can be made
with regards to the prediction of Bjork and Bjork’s (1992) new theory of disuse that spacing practice is not important for relearning:
(i) Although it is impossible to ever "prove" the null hypothesis that there is no difference on long-term performance due to the distribution of practice during relearning, there remains no good evidence to date to argue otherwise.
(ii) If in fact there is an effect of the distribution of practice during relearning, it is probably safe to assume that the effect is fairly small. Thus, the implications of the effect are not as crucial in terms of providing insight as to basic guidelines for ensuring efficient and optimal relearning conditions.

Based on the above two conclusions, the position advocated in this dissertation is that distributing practice during relearning does not make an important difference for long-term performance.

**Practical Implications**

The present findings are likely to be of considerable practical importance in two respects. First, if massed practice is as effective as spaced practice for relearning, then the benefits associated with a massed practice schedule can be obtained without sacrificing the high level of performance typically associated with spaced practice. Second, the results indicate that principles for optimal original learning do not necessarily generalize to relearning.

**Benefits of Massed Practice**

The preliminary finding that massed practice schedules can be as effective for relearning as spaced practice schedules is encouraging for a
variety of reasons, including people's preference for massed practice, as well as the scheduling benefits associated with a massed practice schedule.

Subjective preference for massed practice. In spite of the well-known finding that spaced practice yields better long-term performance than massed practice does, there remains a gap between the scientific findings and application in real-world settings. One possibility as to why this discrepancy remains may have to do with the common misconception that performance during acquisition reliably predicts future performance (Bjork, 1994). During the acquisition or training session, performance with spaced practice is typically inferior to that of massed practice. As a result, people often overestimate the effectiveness of massed practice. Rapid progress in the form of near-perfect performance during massed practice is reassuring to the learner, thereby leading to greater satisfaction and confidence (e.g., see Baddeley & Longman, 1978; Landauer & Ross, 1977; Zechmeister & Shaughnessy, 1980).

The present research suggests that one's current ease of access to information may be a more reliable index during relearning than it is during original learning. The ability for performance from massed practice to be maintained is due, according to the new theory of disuse (Bjork & Bjork, 1992), to the high level of storage strength working to preserve such access, consequently leading to positive long-term effects. As a result, people's metacognitive judgment (i.e., assessment of one's own competence), intuition, and standard practice regarding the effects of distributing practice are not necessarily misguided for relearning. These findings are positive in that the discrepancy between subjective and
objective experience with respect to optimal training conditions may not exist with previously well-learned material.

**Scheduling benefits of massed practice.** In addition to the preference and subjective benefits associated with massed practice are the scheduling and time-efficiency benefits characteristic of a massed practice schedule. The issue of relearning is applicable in many real-world military and industrial settings in what is typically referred to as refresher training. Refresher training can be defined as post-training intervention designed to maintain a high level of performance on a particular skill. In such real-world settings, efficiency of training is often measured in terms of the "cost" of resource expenditures, including "operational costs, student time, course material and equipment, instructors and support personnel" (Hagman & Rose, 1983, p. 200).

In the interest of reducing costs, it is important to try to minimize the time spent in refresher training (Schmidt & Bjork, 1992). Relevant to achieving this goal is the fact that "massed sessions take less total time than do spaced sessions." (Christina & Bjork, 1991, p. 31) The time pressure involved in designing refresher training in terms of using spaced or massed practice is particularly discrepant when the distribution of practice is defined in terms of sessions as opposed to discrete practice trials. Further, Hagman and Rose (1983) note the "scheduling problems and other disruptions associated with the insertion of spaced repetitions within operational training programs" (p. 205).

One study that exemplifies the efficiency of massed sessions in terms of the total number of days to complete training, in addition to the higher
satisfaction associated with such massed training schedules, is that of Baddeley and Longman (1978). In their study, postal workers were taught to type postcodes into a sorting machine. The postal workers were separated into four groups based on their training schedules: learning sessions occurred either once or twice per day, with the duration of the session lasting for either one or two hours. Retention tests were given 1, 3, and 9 months following the training sessions. Although the group that was given training once per day for one hour per day performed the best on retention, this group also took more days to complete the training and rated their satisfaction the lowest because they "felt they were falling behind the groups that were getting more practice per day." (Christina & Bjork, 1991, p. 31)

Original Learning vs. Relearning

The finding of the present research that different principles may apply for relearning compared to original learning is important from a practical perspective, as well as from a theoretical perspective. As Christina and Bjork (1991) state, "An important practical point on refresher training is that the practice or rehearsal needs of retrainees appear to be different from those of new trainees." (p. 36)

The idea that the experience level of the learner might interact with the optimal type of training is not entirely new. In fact, Magill and Hall (1990) raised the question in their review of the contextual interference effect in the motor skill acquisition literature. The contextual interference effect is analogous to the blocked versus random practice effect discussed earlier. Specifically, high contextual interference practice is one involving
random practice conditions and low contextual interference practice is one involving blocked practice conditions. Ironically, Magill and Hall argue for the complete opposite prediction than that supported in the present research. Namely, they propose:

Since high contextual interference practice conditions are more difficult than low contextual interference conditions, it seems reasonable that high contextual interference conditions early in practice may pose a learning problem for beginners and that only after some degree of expertise or prior experience with related skills has been achieved would a high contextual interference practice situation be beneficial.

(pp. 259-260)

Although the above idea was not supported in terms of the preliminary findings from the current research, it is interesting to note the line of reasoning behind their prediction parallels that arguing for an expanding sequence of retrieval practice during original learning (see Landauer & Bjork, 1978). Further, Magill and Hall's notion of a possible interaction between contextual interference and the level of expertise of the learner does appear to be correct.

In addition to the notion that expertise might interact with optimal learning conditions, the current findings are encouraging in that refresher training may be carried out using relatively cost-effective massed scheduling. The current findings add to an on-going list of findings suggesting that "relatively efficient, cost-effective techniques can be used to maintain a given level of original learning or mastery during nonuse periods for retrainees" (Christina & Bjork, 1991, p. 36) Amongst the
techniques Christina and Bjork propose include: (i) using imaginary practice, (ii) expanding the intervals between practice sessions, (iii) reducing the simulation's fidelity, (iv) part-task training, and (v) partial cuing (for further discussion on this topic, see Annett, 1979; and Christina & Bjork, 1991). Now, in addition to these cost-effective techniques, the tentative finding that massed practice is as effective as spaced practice for relearning can be added to the list.

Related Research

Annett (1979) wrote:

Industrial society is entering a period in which any individual's store of skills may not be needed all the time. . . . Expert knowledge of how to provide training that is retained and how to provide appropriate rehearsal and retraining to revive valuable skills when they are required provides a challenge for applied psychologists. (p. 240)

Yet, not much research has been conducted on relearning, or extensive long-term learning, due to inherent limitations within the psychological laboratory. Any such research that has been conducted within the private and military sectors is "rarely reported in the public literature and is difficult to obtain" (Hagman & Rose, 1983, p. 199).

One of the few exceptions to this void in the literature is reported in a paper by Bahrick (1979). Bahrick notes the fact that "memory research has not dealt directly with repeated acquisition processes" (1979, p. 298). Bahrick proceeds to report the results from a study investigating the effect of the time interval separating reacquisition, or refresher, sessions. In the study, the subject's task was to learn the Spanish equivalents to 50 English
words. Bahrick employed a paradigm referred to as the "method of successive relearning." In the first learning session, the 50 English-Spanish paired associates are presented, followed by alternating test and study trials. A dropout technique is used such that only those items failed on the previous test trial are presented on the next study trial. The procedure continues until a single correct response has been made for each of the items. Each of the relearning sessions follows a similar procedure, except that the session begins with a test trial, rather than a study trial.

The intersession interval between the relearning sessions was manipulated such that the sessions were separated by either 0 (i.e., no interval between sessions), 1, or 30 days. A final retention test was administered after either three or six training sessions. Performance during the acquisition trials (as measured by the first test trial on each of the relearning sessions) was clearly superior for the group with an intersession interval of 0 days, followed by the intersession interval of 1 day, and worst for the group with an intersession interval of 30 days. However, consistent with traditional spacing effect findings, on the final retention test administered 30 days after the final reacquisition session, the results reversed, with the 30-day intersession condition clearly performing the best, followed by the 1-day intersession interval, and then the 0-day interval. Bahrick (1979) concluded, "If the access interval is much longer than the retraining interval, much of the acquired information becomes inaccessible during the access interval." (p. 296)

Whereas the study by Bahrick (1979) addressed the question of the optimal spacing of refresher sessions, the present studies address the
question regarding the spacing of the practice within the refresher sessions themselves. Although these two questions are related, they have distinct theoretical differences with regards to the predictions of the new theory of disuse (Bjork & Bjork, 1992). The main difference is in the level of storage strength entering each of the refresher sessions. Namely, by setting a relatively low criterion of one correct trial for the dropout technique employed by Bahrick, the amount of storage strength accumulated within each of the reacquisition sessions would be expected to be fairly low. In contrast, the level of storage strength was required to be relatively high entering the relearning session in the present research. As detailed earlier, the level of storage strength affects the rate at which access to the information is maintained.

**Possibilities for Future Research**

This dissertation represents but a small step in the examination of the effects of distributing practice during relearning. Aside from the obvious need to replicate the current findings and identify boundary conditions, one possibility for future research would be to conduct a better, and even more convincing, comparison experiment. Experiment 1b could feasibly be improved in two ways: First, it would be best if the materials used in the original learning experiment were identical to those of the relearning experiment. This could be done by conducting an analogous study on grade school children who are in the process of learning their mathematical times tables. Second, it would be ideal to be able to show a spacing effect in original learning using the same dependent variable as used in relearning. This second change is more difficult to accomplish due to the
fact that performance on the multiplication problems during original learning would need to be sufficiently high to collect analyzable reaction time data.

The fact that a spacing effect during original learning was not shown using reaction time data in the present study is not all that critical due to the fact that spacing effects have been shown previously when reaction time was the dependent variable. For example, in a study conducted by Shea and Morgan (1979), subjects were taught three different multiple-component arm movement drills. Specifically, the subject's task was to complete each of the movement patterns as quickly as possible, with the dependent variable being the total time (in sec) to complete the drills. All of the movement patterns involved picking up a tennis ball, knocking over a series of wooden barriers, and then returning the ball to a specified location. The movement patterns were learned during acquisition using either blocked or random practice conditions. Under blocked practice conditions, practice on each of the three different movement patterns was practiced sequentially, such that all of the practice trials on a particular pattern were given before moving on to the next pattern (analogous to the massed conditions in the present research). Under random practice conditions, the trials on each of the three movement patterns were randomized such that a particular pattern was never practiced on successive trials (analogous to spaced conditions). A retention test was administered under both blocked and random conditions after either 10 min. or 10 days. At both retention intervals, regardless of whether the retention trials were themselves random or blocked, subjects' performed faster if they had been
given random practice during acquisition.

Thus, the study by Shea and Morgan (1979) demonstrates traditional effects of distributing practice during the initial learning of a task. The fact that the dependent variable in their study was the reaction time to complete a task not only demonstrates the viability of reaction time measures in terms of finding a spacing effect, but also suggests that the paradigm employed in the study might lend itself nicely to be used in future research on relearning. The usage of reaction times as the dependent variable may help alleviate the problem of rapidly attaining asymptotic performance during the relearning of a skill.

To conclude, although much more research is needed to confirm and extend the findings from the present research, the studies reported here provide a preliminary and encouraging step in addressing the question of distributing practice during relearning. The issue of optimal relearning conditions is just one of a list of topics that has been, for the most part, neglected in the literature. As Bahrick (1979) states,

memory research can and should explore the conditions under which acquired information is maintained over long periods. Such questions have great practical and theoretical importance, but psychologists have been unable to deal with them because of constraints imposed by traditional methods." (p. 308)

It is hoped that the research reported here is seen as illustrating not only the theoretical and pragmatic value of this advice, but also a possible means of exploring such questions.
References


Press.


week. Memory & Cognition, 8(6), 528-538.


(Eds.), Practical aspects of memory (pp. 625-632). London: Academic Press.


Behavior, 9, 596-606.
Schendel, J. D., & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. Journal of Applied Psychology,
67(5), 605-610.


Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society, 15*, 41-44.
Appendix


For the following equations, the new theory of disuse (Bjork & Bjork, 1992) predicts that an item (i) in memory has two "strengths:" a storage strength ($S_i$) and a retrieval strength ($R_i$) and that $0 \leq S_i$, $R_i < 1$.

1. $S_{i,n+1} = S_{i,n} + \theta_r (1 - S_{i,n})$
   where $\theta_r = \theta (1 - R_{i,n})$

2. $R_i \rightarrow R_i + \alpha_S (1 - R_i)$
   where $\alpha_S = \alpha S_i$

3. $R_i \rightarrow R_i - \delta_S R_i$
   where $\delta_S = \delta (1 - S_i)$