RECALL AND RECOGNITION TEST EXPECTANCY EFFECTS ON LOW AND HIGH ACHIEVING GRADE 9 AND 10 ADOLESCENTS

by

SUSAN MARIE GOLDSMITH
B.A., UNIVERSITY OF VICTORIA, 1983

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Psychological Foundations in Education

We accept this thesis as conforming

to the required standard

[redacted]

Dr. Beverly A. Timmons

[redacted]

Dr. C. Brian Harvey

[redacted]

Dr. Richard B. May

[redacted]

Dr. Lloyd G. Ollila

University of Victoria

September 1985

Supervisor: Professor Beverly A. Timmons

ABSTRACT

Researchers have found that subjects who expect and receive a recall test perform significantly better than subjects who expect a recognition test and receive an unexpected recall test. This is called the recall test expectancy superiority (TES) effect. Other investigators have found that students of different achievement levels handle organizational, memory, study, and problem-solving tasks quite differently. An experimental study is described in which 36 high achievers and 36 low achievers, at the Grade 9 and 10 level, were led to expect a certain type of test (recall or recognition) for two study-test sessions. Each study-test session required the students to study a one-page story for seven minutes and answer a fifteen-item quiz for five minutes. During the third study-test session, half of the high achievers and half of the low achievers received the opposite type of test from what they expected. The present researcher intended to discover if an unexpected switch to a completion test (when a

multiple choice test was expected) had a different effect on the performance of low and high achievers than an unexpected switch to a multiple choice test when a completion test was expected. Of interest was whether or not high and low achieving students performed differently when expecting completion and multiple choice tests. The results showed that no significant differences were found between the high and low achievers when either group expected one type of test and received the opposite in Session 3. Neither group demonstrated the recall TES effect. Either the high and low achievers did not prepare differently for the different types of tests, or any shifts in their study habits made no difference to their recall performance.

Examiners:

[redacted]

Dr. Beverly A. Timmons

[redacted]

Dr. C. Brian Harvey

[redacted]

Dr. Richard B. May

[redacted]

Dr. Lloyd O. Ollila

TABLE OF CONTENTS

List of Tables

List of Figures

ACKNOWLEDGEMENTS

Financial support for the project was provided through a grant from the Educational Research Institute of B.C. The author wishes to acknowledge the assistance of her committee members, Drs. Timmons, Harvey, and May. Special appreciation is given to her supervisor, Dr. Beverly Timmons, for her support and guidance.

DEDICATION

To my husband Len for his patience,
to my mother and stepfather for their confidence,
to all three, for their love.

Chapter One: Introduction

Testing affects everyone at some point in their lives. It is a vital subject that concerns millions of students and teachers (Frederiksen, 1984). When students are told they will be taking a test, they often ask their teachers what type it will be. Will it be essay, fill-in-the-blank, or multiple choice? Presumably knowing the question types will guide the students' studying behavior. The question is, do students use this information to their best advantage? Past research has established that people who expect a completion (fill-in-the-blank) test and get a completion test perform significantly better than do people who expect a multiple choice test but receive an unexpected completion test instead (Balota & Neely, 1980; Connor, 1977; May & Piekkola, 1983; Schmidt, 1983). This is called the recall test expectancy superiority (TES) effect. The present study investigated whether this effect holds true for both high and low achievers.

Significant differences between high and low achievers have been found in studies that investigated

educational abilities such as organizational (James & Brown, 1973; Rickards & Hatcher, 1977-78; Shuell, 1983; Tyler, Delaney, & Kinnucan, 1983), memory (Kail, 1979; Torgesen, 1977), study (Mikulecky & Olshavsky, 1979; Owings, Petersen, Bransford, Morris, & Stein, 1980; Samuels & Horowitz, 1980; Smiley, Oakley, Wortnen, Campione, & Brown, 1977; Sullivan, 1978), and problem-solving (Kavale & Shriner, 1979; Larkin, McDermott, Simon, & Simon, 1980) skills. Based on these findings of differences, the author's intention was to investigate possible differences between high and low achievers in a test expectancy situation.

The purpose of the study was to investigate the relationship between test expectancy and student achievement. The study focused on the effect of test expectancy with regard to the type of test (completion versus multiple choice) on the performance of high and low achieving students.

Specifically, these were the questions to be answered: do high and low achieving students perform differently when expecting completion and multiple choice tests? Does an unexpected switch to a completion test (when a multiple choice test is

expected) have a different effect on performance of low and high achievers than an unexpected switch to a multiple choice test when a completion test is expected? Please see Appendix A for operational definitions.

A review of the test expectancy research that has occurred over the last fifty years is presented in Chapter Two. Among the test expectancy factors that are discussed are both prose text and word lists as study materials, recall and recognition test expectancy effects, and theoretical explanations of the test expectancy differences found between recall and recognition test performance. As(well, there is a presentation of studies that have found distinguishing characteristics between high and low achievers in organizational, memory, study, and problem solving skills. The author discusses how these differences may apply to a test expectancy situation. In Chapter Three, the rationale of the study is presented, followed by the specific research and null hypotheses of the present research. The methodology of the author's study is also described in Chapter Three. This includes details about the subjects, study and

test materials, treatment administration, and data collection and analyses procedures. The results of the experiment are discussed in Chapter Four. Specific outcomes are considered in relation to the findings of other researchers. In Chapter Five major conclusions and limitations of the study are summarized. Suggestions for future research in the area of test expectancy and achievement are also presented within the final chapter.

Chapter Two: Review of the Literature

Even though there has been interest in classical testing effects for over fifty years (Gersild, 1929), investigations and theoretical explanations regarding test expectancy did not increase noticeably until the early 1970's (Anderson & Bower, 1972, 1974; Kintsch, 1970). Since Underwood (1972) noted the lack of research concerning the effects of test expectancy on both learning processes and the retention of verbal material, various researchers have begun to investigate the area. Most of the early work on test expectancy focused on classroom applications and student performance in field settings (Meyer, 1936; Sax & Collet, 1968). Some of the more recent studies have maintained this focus (Duchastel, 1981); however, several researchers are now investigating the effects of test expectancy on recall and recognition of word lists and prose text (Balota & Neely, 1980; Hay & Pieńkola, 1983).

List Studies

Recall tests.

In list studies using recall tests, the main

finding has been the recall test expectancy superiority (TES) effect. When this effect occurs, people who expect a completion test and get a completion test perform significantly better than do people who expect a multiple choice test but receive an unexpected completion test. When a significant test expectancy is found, it is termed a recall TES effect. This effect is robust and replicable both in list learning (Balota & Neely, 1980; Connor, 1977) and in prose processing (May & Piekkola, 1983; Schmidt, 1983).

Recognition tests.

So far, a significant recognition TES effect has not been reported in the literature on recognition tests. The results concerning a recall or recognition test expectancy on recognition tests have been considerably mixed. Only two studies were found by this writer in which (Hall, Grossman and Elwood, 1976; Neely & Balota, 1981) significance has been reported and that was for a free recall TES. In both of these studies high imagery and high frequency words were used. Conversely, May and Sande (1982) did not find a significant test expectancy by word frequency interaction. In their results rare words were retained

better than common words and subjects expecting and receiving a recall test retained more words (rare and common) than those subjects expecting a recognition test and receiving a recall test. Jacoby's (1973) results indicated a significant cued recall test expectancy, suggesting that cueing may be an important variable. D'Ydewalle's (1981) results showed a non-significant recognition test expectancy effect. The results concerning a recognition TES effect are inconclusive; however, it appears that word-frequency is an important factor in retention.

Other factors investigated in list studies.

Categorized material and interpolated activities may be associated with non-significant trends for recognition test expectancy (Connor, 1977; Jacoby, 1973). Neely and Balota (1981) believed the number of study-test trials was important in inducing a test expectancy effect. Their claim was that more study-test trials would produce a more robust test expectancy than would fewer study-test trials. Interestingly, the test expectancy effect has reached significance with as few as one study-test trial (d'Ydewalle, 1981; Leonard & Whitten, 1983) and as many

as six study-test trials (Balota & Neely, 1980; Neely & Balota, 1981). The number of trials does not seem to be a major factor in inducing test expectancy, at least with college-aged subjects.

Free versus discrete study trials may be an important factor in test expectancy. Out of five experiments using free presentation of the study material (Connor, 1977; Hall, Grossman & Elwood, 1976; May and Sande, 1982), only one showed a significant recall test expectancy effect (Hall, Grossman & Elwood, 1976). The other studies showed a non-significant trend in favor of a recognition TES effect for a recognition test. The study by Hall, Grossman, and Elwood (1976) differed from other researchers' in that it did not use categorized material. In experiments using discrete presentation of study materials, researchers have found a significant free recall TES effect (Neely & Balota, 1981) and a significant cued recall test expectancy effect (Jacoby, 1973). Aside from these studies the results concerning free versus discrete study trials have been mixed (Connor, 1977; d'Ydewalle, 1981; Hall, Grossman, & Elwood, 1976, Experiment 3).

Encoding and retrieval processes were used in theoretical explanations of recall and recognition expectancies (Anderson & Bower, 1972; Kintsch, 1970). One problem with the theoretical explanations, which will be discussed later, is that they were based on post hoc methods of analysis. Because the data collection paradigms did not investigate the encoding and retrieval strategies while the subjects were studying, researchers have examined serial position effects in recall performance (d'Ydewalle, 1981; May & Sande, 1982). Serial position effects refer to one part (beginning, middle, or end) of a study passage being recalled better or worse than another part of a study passage. May and Sande (1982) found that serial position curves were very different for recall versus recognition expectancies. The recall expectancy position curves showed a strong primacy effect (words from the beginning of a list were remembered better than words from the end of the list) and a linear slope whereas recognition expectancy position curves showed no primacy effect and a flattened serial position curve. From this, May and Sande (1982) proposed that subjects expecting a recognition test will use

different study strategies than will subjects expecting a recall test.

Prose Studies

Recall tests.

Most of the studies investigating test expectancy have concentrated on list learning while there has been a limited number of prose studies reported (d'Ydewalle & Roselle, 1978; d'Ydewalle, Swerts, & De Corte, 1983; May & Piekkola, 1983; Schmidt, 1983).

As in list studies, the cued recall TES effect in prose studies is robust (d'Ydewalle, Swerts, & De Corte, 1983; May & Piekkola, 1983; Schmidt, 1983). In other words people who expect a completion test perform significantly better than do people who expect a recognition test when both groups are given a cued recall test. There is evidence, however, which indicates both that the recall TES effect may be limited to cued recall in prose studies and that it is less detectable in free recall prose tasks (d'Ydewalle, Degryse, & Swerts, 1982; Rickards & Friedman, 1978). In comparing the completion performance of subjects expecting a completion test to subjects expecting a recognition test, D'Ydewalle et al., (1982) and

Richards and Friedman, (1978) found no significant differences. This outcome is important because free recall tasks were used in both of the studies and no recall TES effect was found. It seems that the recall TES effect may, indeed, be limited to cued recall tasks. Studies using prose text and cued recall tasks have found recall expectancy advantages (d'Ydewalle, Swerts, & De Corte, 1983; May & Piekkola, 1983). The type of recall demanded (prose "gist" - free recall versus question answering - cued recall) seems to have a direct effect on the strength of the recall TES effect.

Schmidt (1983) found that organization of material (unrelated versus related sentences) and type of material (fiction versus nonfiction) did not interact with test expectancy significantly. In the same experiment he also found that people remembered fictional material better than nonfictional material even though the sentences were unrelated.

As in the list studies, the number of study-test trials, ranging from none (Schmidt, 1983) to four (May & Piekkola, 1983) did not seem to effect the robustness of the recall TES effect in prose. The same result was

found when interpolated activities (Hay & Piekkola, 1983), and varying study periods (d'Ydewalle, Swerts, & De Corte, 1983) were used. Free study formats (d'Ydewalle & Rosselle, 1978) versus discrete study formats (Schmidt, 1983) did not influence the significant recall TES effect.

Recognition tests.

Several studies indicated a cued recall TES effect (d'Ydewalle, Swerts, & De Corte, 1983, Experiment 2; Hay & Piekkola, 1983; Schmidt, 1983) in recognition tests (except d'Ydewalle & Rosselle, 1978). This effect held up under different numbers of study-test trials and free or discrete study formats (Hay & Piekkola, 1983; Schmidt, 1983). Free study formats may have facilitated the greatest cued recall TES effect (d'Ydewalle, Swerts, & De Corte, 1983). Interpolated activities, such as counting backwards by threes (Hay & Piekkola, 1983), did not seem to alter overall results. Jacoby (1973) found the same effect in his list study data.

Theoretical explanations

Some of the theoretical explanations of test expectancy have been based on encoding and retrieval

processes. Underwood (1972) suggested that recognition and recall encoding were qualitatively different, each being affected differently by test expectancy. In explaining the differences between recall and recognition processes, Leonard and Whitten (1983) suggested that recall and recognition processing involved different types of association. Recall associations may have been interitem (item-to-item), and recognition associations may be contextual (item-to-context). It also was proposed that the recall TES effect was due to some sort of deeper processing than was present in recognition tasks (Craik & Lockhart, 1972).

Anderson and Bower (1972, 1974) refered to the generation of associative networks (pathway coding) and a matching process called node-tagging. They speculated that recall processes involved pathway coding, while recognition processes involved only part of the recall process which has been called occurrence, context, or node-tagging. In explaining recognition encoding, Anderson and Bower claimed a tag was created between the entire experimental context and each word at the moment of presentation. Once a whole list of

words was presented, the person would have tags for the words which were called list markers. The probability of correct recall depended on the adequacy of the list marker formation during the study trial. Recall encoding depended on both pathway coding and node-tagging. Each list word had many associative pathway extensions. Once a list marker was formed, only one of the possible pathways was used to search for an associated word in the list. The list marker of the associated word was tagged to the original word's list marker if it was found. Recall processes were thought to generate more associative networks and fewer node-tags, and recognition processes were thought to generate more node-tags and less associative networks. Kintsch (1970) has supported the pathway coding - node-tagging theory also.

Rickards and Friedman (1978) found that the quality of performance may vary depending on the type of test expected. In their study the subjects expecting an essay (recall) test chose to take notes on sentences of greater importance to the overall meaning of the passage; those subjects expecting a multiple choice (recognition) test did not take notes. An

explanation by Hall, Grossman, & Elwood (1976) suggested that the amount or type of effort may be important. They claimed the recall TES effect may have been due to the subjects' perceiving the recall task as being more difficult than the recognition task. If this is so, subjects may apply more concentrated effort on the recall task.

Gillund and Shiffrin (1984) proposed an extension of the Search of Associative Memory (SAM) model (Raaijmakers & Shiffrin, 1981) for recall. The extension of the model was based on the assumption that a familiarity process was used for recognition. Their computer simulation model suggested that increases in the interitem parameter and the self-coding (rote repetition) parameter improved recognition performance by increasing the familiarity of the target distribution, but not the distractor distribution. Increases in interitem parameters also aided free recall by improving recovery and reducing recall self-sampling when item cues were used; however, increases in the self-coding parameter harmed free recall by increasing recall self-sampling. The authors defined self-sampling as the tendency for an item that

was used as a cue to sample its own image.

Gillund and Shiffrin offered the only model found by this writer that attempted to deal explicitly with test expectancy effects. They suggested increased expectancy for recall tests results in increased interitem coding, which may have in turn decreased self-coding. According to their model, both recall and recognition performance have been predicted to rise from increased interitem coding. They cited several studies which indicated varying the test expectancy from recall to recognition increased both recall and recognition performance (Balota & Neely, 1980; Hall, Grossman, & Elwood, 1976; Neely & Balota, 1981).

There are no conclusive explanations of exactly how test expectancy affects recall and recognition performance (Schmidt, 1983). As mentioned earlier, most explanations are based on post hoc analyses. Since procedures such as serial position analyses (May & Sande, 1982) have not led to any firm conclusions, direct observation of performance (ie. eye movement activity) may be necessary to the understanding of encoding and retrieval strategies that occur while subjects are studying for recall and recognition tests.

High and Low Achievement

There do not appear to be any studies in the test expectancy literature which compare high and low achievers; however, achievement is investigated in association with other educational abilities such as organizational, memory, and problem solving skills. In relation to test expectancy both organization of study material (Jacoby, 1973; Schmidt, 1983) and memory (Anderson & Bower, 1972, 1974; Gillund & Shiffrin, 1984) have been investigated but without considering the achievement factor. The following studies provide evidence which indicates high and low achievers perform differently in organizational, memory, and problem solving tasks. Because distinct differences have been found between high and low achievers in these educational skills, the author proposes that distinct differences between the two achievement groups may be found in a test expectancy situation. The review focuses mainly on studies investigating reading performance since the experimental task of the present study is based on reading and comprehension. The studies on problem-solving have been included in the review because they highlight the ways in which

subjects of different abilities approach learning tasks. The present study compares how high and low achievers perform on multiple choice and completion tests when a certain test type is expected. High and low achievers are defined operationally by the researcher(s) of each study. The specific definition of high and low achievement used by each author is not as important to this review as are the distinct differences which were found between subjects of different achievement levels.

Organizational skills.

Shuell (1983) investigated whether or not instructing students to organize recall would differentially facilitate the performance of "fast" and "slow" learners. In Experiment 1 he asked the learners to organize their list word recall alphabetically. The number of words recalled by both the fast and slow learners increased significantly. Shuell proposed that the type of organization suggested was either overlooked or not normally used by the learners. In Experiment 2 the learners were asked to organize their word recall categorically. The results were facilitated performance for the slow learners and

little difference in performance for the fast learners. When the learners were informed of the organization of their study materials and encouraged to use appropriate study strategies, the difference in performance between the fast and slow learners was cut in half. From this, Shuell concluded that the difference between fast and slow learners may lie in their ability to use organizational strategies most appropriate for the material being learned. Because the organizational instructions had a facilitative effect on the performance of only the slow learners, Shuell proposed that the learning strategy was more effective than the one normally used by slow learners and either less effective or, at least, no more effective than the one used by average and fast learners.

Tyler, Delaney, and Kinnucan (1983) investigated good and poor readers' abilities to use advance organizers, which the authors described as a short paragraph used to bridge the gap between what the student's already knew and information they were about to learn. The subjects were asked to read short passages preceded by one of two types of advance organizers; one was a summary of the passage, and the

other contained definitions and explanations of major concepts needed to understand the article. The second type of advance organizer followed Ausubel's (1968) criteria of relating the information to what the reader should already know. The researchers found that good readers showed greater recall of detail given either type of advance organizer, while poor readers displayed enhanced recall of detail for the Ausubel-type advance organizer and not the summary organizer. The authors suggested that the Ausubel-type organizer may have compensated for the poor readers' organizational deficits.

Rickards and Hatcher (1977, 1978) studied the effect of interspersing meaningful learning questions in text prose as semantic cues for poor comprehenders. They had the subjects read 800-word passages, and, after every two paragraphs, they received a meaningful learning question requiring the organization of facts under given concepts, a rote learning question requiring literal recall of passage information, or no learning question of any kind. For good comprehenders the type of treatment did not have a significant effect on text recall; however, for poor comprehenders the

meaningful learning questions facilitated significantly greater recall of facts than did the rote learning questions or no question at all. Rickards and Hatcher claimed their results supported the usefulness of Ausubel's theory (1968) in constructing questions appropriate for poor comprehenders. It appeared that poor comprehenders did not spontaneously organize paragraph details around main ideas of a passage. Conversely, good comprehenders seemed to generate a context while reading text, thereby decreasing the need for context-providing cues.

James and Brown (1973) studied the effects of prose organization and individual differences on free recall. They investigated three main topics: (a) the effect of paragraph organization on free recall of sentences and on the selection of clustering strategies; (b) the effect of different organizational skills on a person's recall and strategy selection; and (c) the effect of organizational skills on other cognitive abilities. The students were asked to study paragraphs organized by concept names, concept attributes, or at random. The group that studied paragraphs organized by concept names recalled the most

correct statements, although clustering by names was predominant for all groups. A relationship between cognitive factors and recall scores was found for each group. More importantly, superior recall emerged when the students learned a highly organized passage and a preferred recall strategy. Highly organized students were not greatly influenced by the organization of the paragraphs, whereas low organizers were. The results indicated that students low in organizational skills may have required highly structured materials, while students high in organizational skills performed similarly on materials that were high and low in structure.

These studies point out the importance of providing low achievers with specific instuctions (Shuell, 1983), appropriate advance organizers (Tyler et al., 1983), meaningful learning questions (Rickards & Hatcher, 1977-78), and highly structured materials (James & Brown, 1973). If these results can be generalized to a test expectancy situation, low achievers may require instructions to use specific study strategies for recall versus recognition tasks. The organizational data indicates that low achievers

might not spontaneously select and employ different strategies for different types of tests which could result in lower recall performance.

Memory skills.

Torgesen (1977) investigated the short-term memory (STM) abilities of good and poor fourth grade readers. The subjects were asked to study and memorize various materials while being discretely observed by the experimenter. The children in the two reading groups differed significantly both in their recall scores and in the different types of study behavior. Torgesen reported that the good readers approached the reading tasks in an active and organized manner, while the poor readers reflected a lack of ability or inclination to use efficient memorization strategies. Moreover, when both groups of children were instructed in the use of efficient strategies, the poor readers performed as well as the good readers.

In studying individual differences in children's memory, Kail (1979) found that eight-year-old children were much more likely to use strategy-free methods than strategy-based methods to recall word lists and pictures. The strategy-free methods the eight-year-old

children employed included recency recall, uncued recency judgments, and cued recall. In comparison, eleven-year-old children in the study used strategy-based techniques such as primacy recall, cued recency judgments and free recall to aid recall. Kail suggested that general strategic ability may be a source of individual differences, related to development, in childrens' ability to recall information.

Torgesan's study on STM in fourth grade readers illuminates the importance of providing low achievers in reading with specific instructions in memory strategies. These findings may also suggest that unlike high achievers, low achievers could need specific instructions for the times they are preparing and expecting recall versus recognition tests.

Study skills.

Owings, Petersen, Bransford, Morris, and Stein (1980) investigated the spontaneous monitoring and regulation of learning in successful and less successful fifth graders. The students were asked to read and study paragraphs that were either sensible and straighforward or less sensible and not logical. For

both the successful and less successful students cued recall test scores were much higher for stories that made sense than for stories that did not. The successful students spontaneously monitored their learning as they read and studied, choosing to study difficult stories longer than easy stories. They were cognizant of difficulty in learning the less sensible stories, and could explain why they were having difficulty. The less successful students did not rate the difficulty level of the less sensible stories accurately. They could distinguish between the two types of stories, but only did so when they were prompted. When allowed unlimited study time, the less successful students studied equal times for the sensible and less sensible stories. The authors suggested that many students may have performed below their ability because they did not spontaneously monitor and regulate their learning.

When Sullivan (1978) compared the study strategies of good and poor comprehenders she found that there was a difference in the matching procedures used when they processed factual and conclusive statements. In evaluating factual statements, good readers looked for

consistency by continually matching words in statements with words of similar meaning in the passages. Most of the poor readers were more literal, focusing on nouns and verbs in isolation rather than word clusters. The poor readers found it difficult both to transpose information and to modify literal statements into meaningful thougnt patterns. When the poor readers tried to transpose the passages, they focused on nouns and verbs rather than encapsulation of general ideas. Moreover, poor readers had more difficulty making judgments about conclusive statements, finding supportive examples, and using matching procedures. Seventy percent of the poor readers had difficulty relating past knowledge to the reading material. Sullivan suggested that more structured, organized teaching approaches be used to facilitate strategies in poor comprehenders. Her practical examples included instruction in paraphrasing passages rather than literal recall, sifting out meaningful word clusters for use in deciding between alternatives, searching for supportive examples, and making deductions by combining ideas.

With regard to literal recall Samuels and Horowitz

(1980) found similar results to those of Sullivan (1978). Their sixth grade subjects were exposed to expository text in both oral and written modes. When asked to retell what they had either heard or read, the results from a multiple choice test revealed significant differences in recall from the oral mode did not exist between good and poor readers; in fact, the poor readers performed better than did the good readers in the free recall of literal ideas from the oral mode of presentation. The good readers showed superior recall in the reading mode of presentation.

In comparing the study strategies of good and poor university level readers Nikulecky and Olshavsky (1979) found few significant differences between the two groups. The two main differences found between good and poor readers included the good readers both taking a more serious systematic approach to studying textbooks and reading for more of an overview than did the poor readers.

Smiley, Oakley, Worthen, Campione, and Brown (1977) tested the comprehension ability of adolescent good and poor readers. They measured the recall of thematically relevant material from written prose

passages and from passages presented orally. Good readers were able to recall a greater proportion of the stories in both the reading and listening conditions. Their recall of a particular unit was based on the importance of the unit in the story. Overall, poor readers recalled less of the stories in both the reading and listening modes, and their recall was not related to unit importance within the story. The authors suggested that poor readers had a comprehension deficit which was general in nature, affecting both reading and listening comprehension.

These findings concerning study skills delineate some of the differences between high and low achievers. Less successful students are less likely both to monitor their learning (Owings et al., 1980) and to respond to less structured teaching approaches (Sullivan, 1978) than are more successful students. Researchers have also found that poor readers do not base recall on key ideas within a story (Smiley et al., 1977). These studies stimulate questions concerning the existence of differences between high and low achievers when they study for and expect different types of tests.

Problem-solving and reasoning strategies.

A study by Kavale and Schreiner (1979) identified reasoning strategies used by above-average and average readers in responding to reading comprehension questions. Both groups used similar strategies but above-average readers applied them more often and with more success than the average readers. The above-average readers seemed to be active in their search for understanding and were more likely to use alternative reasoning strategies for different question types than the average readers. This showed a flexibility in the above-average readers' ability to think and reason verbally that was not demonstrated by the average readers. The average readers used similar strategies to the above-average readers but they applied them less often and with less success. The average readers were not as able to select useful alternatives to answering the questions as were the above-average readers.

In comparing expert and novice performance in solving physics problems Larkin, McDermott, Simon, and Simon (1980) found several differences in solution processes. The expert had both strong mathematical

skills and extensive experience in solving problems in mechanics. The expert solved the problems in less than one-quarter of the time required by the novice, and the expert made fewer errors. Secondly, the novice solved most of the problems by working backward from the solution, while the expert usually worked forward toward the solution. A third difference found between the experts and novices was that the novices mentioned each equation that they were about to use and then substituted the values of the independent variables. The expert usually mentioned only the numerical result of the substitution, and not the equation. The authors suggested both that the expert had a direct store of the entire procedure and that the novice stored only the particular equations and their potential use. The novices employed several tests before arriving at the correct answers, whereas the experts' performance seemed automated and answers were reached with very few trials. Larkin et al. claimed that their results indicated that different types of cognitive processing occurred for expert versus novice problem solvers.

From their research findings, Kavale and Shreiner (1979) and Larkin et al. (1980) suggested that students

of different skill levels approach learning tasks quite differently which may result in different test performance. This author intends to discover if recall performance differences exist between high and low achievers on multiple choice versus completion tests.

Chapter Three: Statement Of The Problem And Experimental Proceaures

Rationale

Very few studies have investigated test expectancy effects either in children (Hall, Miskiewicz & Murray, 1977; Maus, Ornstein, & Kreshtool, 1977) or adolescents (May & Sande, 1982). Most studies have used word lists rather than prose as the study material. Some studies have compared completion to multiple choice effects within one group of subjects. This has led to the discovery of the robust recall TES effect. There appear to be no studies in the literature which compare retention means within and between subgroups of high and low achieving adolescents. In this study, the completion and multiple choice tests means of these two achievement groups will be compared.

Hypotheses

In the following hypotheses all of the completion items are cued unless otherwise specified.

1. On the final test, within the completion group, the mean test score of those expecting and receiving a completion test will be significantly greater than

the mean test score of those expecting a multiple choice test and receiving a completion test.

2. On the final test, within a completion group, the mean test score of high achievers expecting and receiving a completion test will not be significantly different from the mean test ssore of high achievers expecting a multiple choice test and receiving a completion test.
3. On the final test, within the completion group, the mean test score of low achievers expecting and receiving a completion test will not be significantly different from the mean test score of low achievers expecting a multiple choice test and receiving a completion test.
4. On the final test, within the multiple choice group, the mean test score of those expecting and receiving a multiple choice test will not be significantly different from the mean test score of those expecting a completion test and receiving a multiple choice test.
5. On the final test, within the multiple choice group, the mean test score of high achievers expecting and receiving a multiple choice test

will not be significantly different from the mean test score of high achievers expecting a completion test and receiving a multiple choice test.

6. On the final test, within the multiple choice group, the mean test score of low achievers expecting and receiving a multiple choice test will not be significantly different from the mean test score of low achievers expecting a completion test and receiving a multiple choice test.

Subjects

The subjects were drawn from Cedar Hill Junior Secondary School in the Greater Victoria School District No. 61. Sixty-seven Grade 10 and twenty-three Grade 9 students were selected by the school counsellor as low and high achievers. Thirty-five of the low achievers and thirty-nine of the high achievers returned parental consent forms and were present on the day of testing. To end up with nine subjects in each of the eight conditions the researcher returned the following day to test the low achiever who had been absent and randomly chose three of the high achievers' data and discarded their results. Thirty-six high

achievers (16 Grade 10 females; 8 Grade 10 males; 5 Grade 9 females; 7 Grade 9 males) and thirty-six low achievers (12 Grade 10 females; 16 Grade 10 males; 4 Grade 9 females; 4 Grade 9 males) who participated were considered the final sample. The adolescents were fifteen to sixteen years of age.

The selection of high and low achievers was carried out by the school counsellor, and it was based on the students' letter grades in their English classes. At the time of selection, the students had received three of their four interim report cards. The researcher's criterion for a high achiever was a student's receiving a "B" or higher in at least two of the three English class grades. To qualify as a low achiever the student had to receive a "C-" or lower in at least two of the three English class grades. The grading scale was as follows: A, B, C+, C, C-, D, E, and F. The selection of subjects was non-random; however, assignment of subjects to treatment groups was done randomly using random numbers generated by a computer. Parental permission was required before participation in the study. (See Appendix B for parental consent forms).

Instrumentation

Study Materials. The materials used in sessions one through three were stories taken from Science Research Associates (SRA) materials (Goddard, 1985). They were designed for use in grade 5 to 7 classes. Each story contained from 492 to 500 words and had a readability level between grades 5.5 and 7.6 (Fry, 1977).

Test Materials. The test materials consisted of 15 knowledge level questions per story (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). The questions were not presented in a random order and not the same order as they appeared in the study materials. Parallel forms of the test items were generated in both multiple choice and completion formats (see Appendix C for the study and test materials).

Treatment Administration

Once the low and high achieving students were identified by the researcher's criteria, they were randomly assigned to treatment groups. The names of the high and low achievers were numbered from 1 to 39 and 1 to 36 respectively. Using random numbers from a computer program, 19 students were assigned to multiple

choice questions and 20 students were assigned to completion questions in the high achievement group. In the low achievement group 18 students were assigned to the multiple choice questions, and 18 students were assigned to completion questions. This assignment remained the same for sessions one and two. In the third session, the test type was changed for half of the subjects in each achievement group. Ten high achievers and nine low achievers were randomly selected from the multiple choice test group to take the completion test questions. Similarly, nine high achievers and nine low achievers were randomly selected from the completion test group to take the multiple choice test questions (see Figure 1).

The order of story presentation was counterbalanced across all subjects using the Latin square counterbalancing technique. The three stories were presented in one of the following orders: ABC, BCA, or CAB. Since there were three possible orders, the subject numbers in each treatment group were in multiples of three. The counterbalancing controlled for possible order effects.

One story was presented per session. There were

Figure 1. Research design.

	Session 1	Session 2	Session 3
High Achievers	Multiple Choice (Recognition) n=18	Multiple Choice (Recognition) n=18	Multiple Choice (Recognition) n=9 (1)*
			Completion (Recall) n=9 (2)*
	Completion (Recall) n=18	Completion (Recall) n=18	Multiple Choice (Recognition) n=9 (3)*
			Completion (Recall) n=9 (4)*
Low Achievers	Multiple Choice (Recognition) n=18	Multiple Choice (Recognition) n=18	Multiple Choice (Recognition) n=9 (5)*
			Completion (Recall) n=9 (6)*
	Completion (Recall) n=18	Completion (Recall) n=18	Multiple Choice (Recognition) n=9 (7)*
			Completion (Recall) n=9 (8)*

N = 72

* Refers to condition code group number.

Factors:

1. Type of practice: Recall, Recognition
2. Achievement: High, Low (trait variable)
3. Final task: Same, Different
4. Order of stories: ABC, BCA, CAB

Dependant Variable:

1. Retention score (percent correct)

2 x 2 x 2 x 3 factorial design
(Type of Practice x Achievement x Final Task x Order)

three stories and three sessions altogether. The three sessions were run consecutively in one day over a period of one and one half hours which is a usual procedure in test expectancy experiments (Hall et al., 1976; May & Sande, 1982).

Two experimenters conducted the experiment. One of them, the author, was female and her assistant was male. They divided the procedural duties evenly and attempted to appear equally in charge during the experiment. The author felt the presence of both a male and a female experimenter would, perhaps, combat experimenter bias due to gender. It was believed that this was an important concern, given the age of the subjects (Wrightsman & Deaux, 1981).

Data Collection And Classroom Procedure

Each subject received two large brown envelopes. In the envelope on the left side of each desk were the three stories and three tests, each with a cover page and all in the exact order of the procedure. The other envelope was placed on the right side, of each desk and served to collect the finished tests. The two experimenters introduced themselves saying they were interested in how adolescents read and how they answer

questions about stories (See instructions used in Appendix D). They were told they would receive three stories and be asked to answer a fifteen item quiz about each story. The students were asked to study the way they normally would, and they were told that they could put pencil marks on the study materials.

Analysis Of The Data

In scoring the subjects' protocols one point was awarded for each correct response. For the multiple choice tests the correct response had to be circled for the subject to be awarded a point. Alternatively, for the completion tests the subjects were required to give the same response that would have been circled on the multiple choice version of the test to receive credit. This was considered appropriate since the multiple choice and completion tests for each story were parallel forms. The reliability of each of the tests was measured by the Kuder-Richardson 20 (K-R 20) formula.

A two-way analysis of variance was carried out on the order (ABC, BCA, or CAB) and achievement group (high and low achievers) variables for the two types of practice (recall or recognition), for Sessions 1 and 2.

The results from Session 3 were analysed by an 8 (Group) X 3 (Order) two-way ANOVA. Additionally, all of the completion scores from Session 3 and all of the multiple choice scores from Session 3 were analyzed separately. Two 2 (Final Task: Same, Different) X 2 (Achievement: High, Low) ANOVA'S were conducted, one for the completion data and the other for the multiple choice data.

Chapter Four: Results and Discussion

Results

The means and standard deviations for low and high achievers are presented in Table 1. All of the values are based on raw scores out of fifteen for the subjects in each condition group. Figure 1 details the meaning of each condition number. Within each achievement group the multiple choice means were consistently, although not significantly, higher than the completion means. In comparable conditions, the high achievers repeatedly attained higher scores than did the low achievers. The standard deviations for the completion tests were greater than those for the multiple choice tests for Session 1 and Session 2. For Session 3 where half of the subjects received the unexpected switch, the degree of dispersion was less for the subjects who were not surprised.

Reliability.

The internal consistency of each of the three tests was measured by the Kuder-Richardson (K-R 20) formula. Although the study and test periods did have time limits, none of the subjects appeared to be short

Table 1

Means And Standard Deviations (SD) For Low And High Achievers In Each Condition

	Session 1		Session 2		Session 3	
condition	Mean	SD	Mean	SD	Mean	SD
1	10.42	2.71	10.92	2.28	11.33	2.57
2	11.92	1.93	11.58	1.56	8.67	3.11
3	7.83	2.86	8.92	2.54	10.25	2.63
4	8.92	2.31	8.83	1.90	9.58	1.17
5	8.89	3.44	9.11	1.76	8.89	2.71
6	8.22	2.95	8.89	2.89	5.22	3.67
7	6.33	3.57	6.56	3.58	7.78	4.52
8	5.89	3.33	5.11	5.23	3.67	4.33

of time. This was evident during the experimental sessions and in the scoring of the test protocols. The reliability coefficients ranged from r=.61 to r=.83 on both versions of the three tests. In general, the completion tests appeared to be more reliable than the multiple choice tests. As shown in Table 2, the completion tests resulted in a somewhat greater degree of score dispersion than the multiple choice tests for two of the three tests.

Order, Achievement Group, and Practice.

The story order (ABC, BCA, or CAB) and achievement group (high, low achievers) variables were analyzed for Sessions 1 and 2 to find any significant differences that existed between the high and low achievers. Separate analyses were carried out for recall and recognition scores over the first two sessions. In each case the overall analysis was a 2 (achievement group) X 3 (orders) ANOVA with repeated measures only on the sessions factor. A 2 X 3 (Achievement Group: High, Low Achievers X Order: ABC, BCA, or CAB) analysis of variance on the recognition practice Sessions 1 and 2 revealed significant main effects of order for Session 2, F (2,36) = 4.76, p <.02, achievement group

Table 2

Reliability Coefficients (K-R 20) And Standard Deviations (SD) Of Each Test Including All Subjects For Sessions 1 And 2

	Test type and SD			
Story	Multiple Choice	SD	Completion	SD
Little	.76	3.19	.65	2.77
Land	.62	2.69	.83	3.51
Magic	.61	2.73	.79	3.48

for Session 1, F (1,36) = 9.33, p <.005, and achievement group for Session 2, F (1,36) = 12.30, p <.005. In each of the three orders of presentation, the high achievers consistently attained higher recall scores than the low achievers in both Sessions 1 and 2. The high achievers' scores increased from a mean of 11.17 for Session 1 to 11.25 for Session 2 and the low achievers' scores increased from a mean of 8.49 for Session 1 to 9.15 for Session 2. The overall recall mean for both high and low achievers increased only slightly from 10.05 for Session 1 to 10.28 for Session 2. The analyses did not reveal a significant achievement group by order interaction.

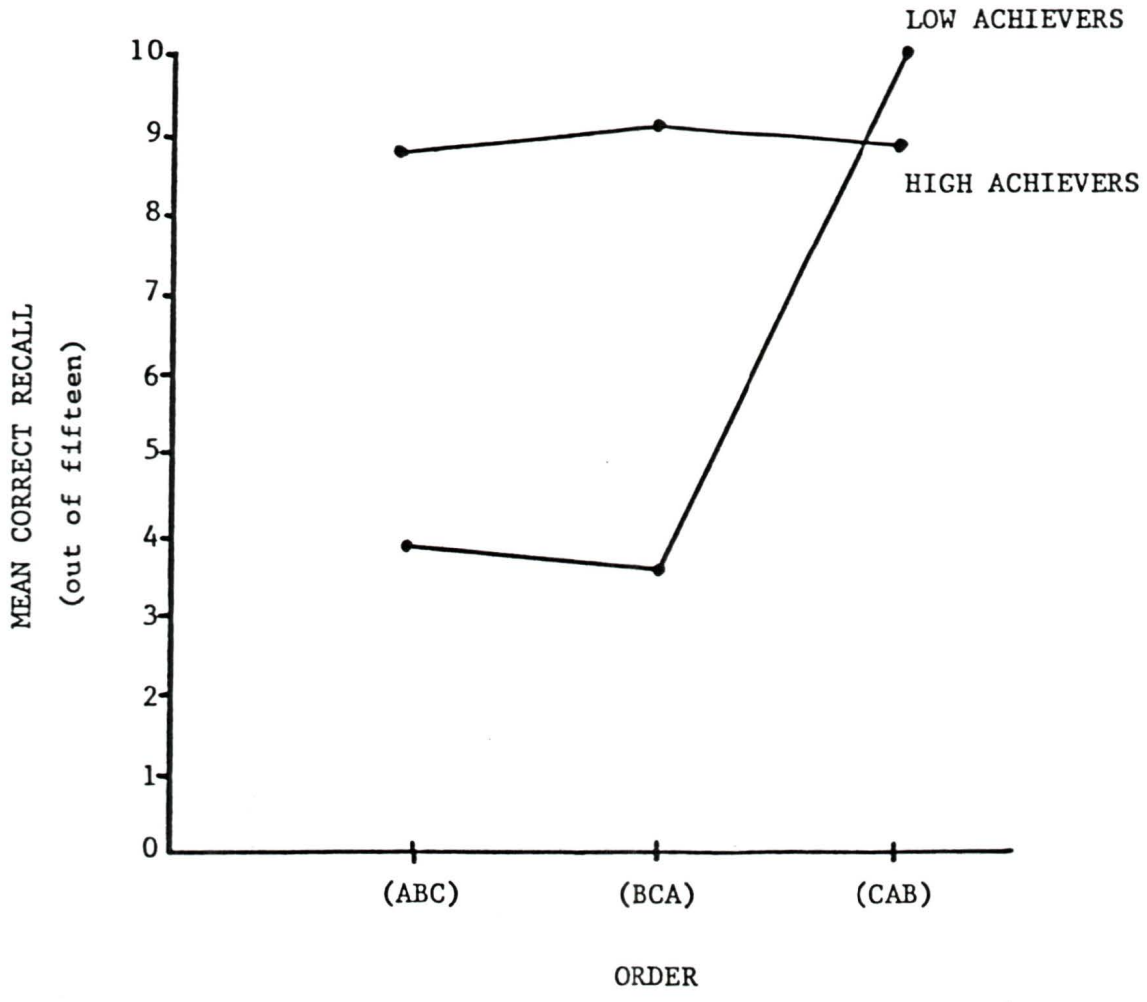
A 2 X 3 (Achievement Group: High, Low Achievers X Order: ABC, BCA, or CAB) analysis of variance for the recall practice Sessions 1 and 2 revealed significant main effects of order for Session 1, F (2,36) = 4.07, p <.05, and for Session 2, F (2,36) = 5.69, p <.05 and of achievement group for Session 1, F (1,36) = 8.42, p <.05, and for achievement group for Session 2, F (1,36) = 11.82, p <.05. As in the recognition results, the high achievers consistently outscored the low achievers in their recall for both Session 1 and

Session 2. As shown in Figure 2 there was one exception to this consistent pattern in Session 2 where the low achievers attained higher recall scores for the CAB order combination. There was a difference between the scores from Session 1 to Session 2 for both the high achievers ($\bar{X}$ = 8.38 to $\bar{X}$ = 8.88) and the low achievers ($\bar{X}$ = 5.89 to $\bar{X}$ = 5.79). The high achievers' recall scores increased from Session 1 to Session 2, while the low achievers' recall scores decreased from Session 1 to Session 2. For the entire sample, the mean increased from 7.40 for Session 1 to 7.57 for Session 2. The analyses revealed a significant Achievement Group X Order interaction for Session 2, F (2,36) = 5.64, p <.05. Figure 2 illustrates the effect that different order combinations had on the recall of the high and low achievers. The order combinations among the high achieving subjects seemed to have a minimal effect on their recall performance. Conversely, the order combination of CAB seemed to give those low achieving subjects a distinct advantage in their recall performance.

Final Task.

In analyzing the results from Session 3, an 8

Figure 2. Mean score correct recall (completion items) as a function of order in Session 2.



(Group) X 3 (Order) analysis of variance revealed a significant main effect due to group, F (2,60) = 6.70, p <.001. There was not a main effect due to order nor an interaction of group X order. The results were further analyzed to address the major research hypotheses. All of the completion scores from Session 3 and all of the multiple choice scores from Session 3 were analyzed separately. See Figures 3 and 4.

In order to make the above comparisons, two 2 (Final Task: Same, Different) X 2 (Achievement: High, Low) ANOVA'S were conducted, one for the completion data and the other for the multiple choice data, as in Figures 3 and 4. The analyses for the completion data revealed a significant main effect due to achievement in the completion data, F (1,38) = 22.58, p <.001. Because there was neither another significant main effect nor a significant interaction found the completion means, research Hypothesis 1 was not accepted:

1. On the final test, within the completion group, the mean test score of those expecting and receiving a completion test will be significantly greater than the mean test score of those

Figure 3. Mean score correct (out of fifteen) for completion scores.

		Receive Completion (Recall)		
		High Achievers	Low Achievers	
Expecting	Completion (Recall)	$\bar{X}$=9.58 (Group 4)	$\bar{X}$=3.67 (Group 8)	Total $\bar{X}$=6.63
	Multiple Choice (Recognition)	$\bar{X}$=8.67 (Group 2)	$\bar{X}$=5.22 (Group 6)	Total $\bar{X}$=6.95
	Total	$\bar{X}$=9.13 *	$\bar{X}$=4.45 *	

* $p < .001$

Figure 4. Mean score correct (out of fifteen) for multiple choice scores.

		Receive Multiple Choice (Recognition)		
		High Achievers	Low Achievers	
Expecting	Completion (Recall)	$\bar{X}$=10.25 (Group 3)	$\bar{X}$=7.78 (Group 7)	Total $\bar{X}$=9.02
	Multiple Choice (Recognition)	$\bar{X}$=11.33 (Group 1)	$\bar{X}$=8.89 (Group 5)	Total $\bar{X}$=10.11
	Total	$\bar{X}$=10.79 *	$\bar{X}$=8.34 *	

* $p < .05$

expecting a multiple choice test and receiving a completion test.

and null Hypotheses 2 and 3 were not rejected:

2. On the final test, within a completion group, the mean test score of high achievers expecting and receiving a completion test will not be significantly different from the mean test score of high achievers expecting a multiple choice test and receiving a completion test.
3. On the final test, within the completion group, the mean test score of low achievers expecting and receiving a completion test will not be significantly different from the mean test score of low achievers expecting a multiple choice test and receiving a completion test.

The ANOVA on the multiple choice data also revealed a significant main effect due to achievement, F (1,38) = 6.36, p <.02. As in the completion data, no other significant differences were found in the multiple choice data. Interestingly, the difference in recall performance between the high achievers ($\bar{X}$ = 10.79) and the low achievers ($\bar{X}$ = 8.34) was significant (p <.05) on the multiple choice tests; however, the

difference in recall performance on the completion tests was even more significant (high achievers: $\bar{X}$ = 9.13, low achievers: $\bar{X}$ = 4.45, $p < .001$).

The lack of any other significant main effects (besides achievement) and interactions resulted in failure to reject null Hypotheses 4 through 6:

4. On the final test, within the multiple choice group, the mean test score of those expecting and receiving a multiple choice test will not be significantly different from the mean test score of those expecting a completion test and receiving a multiple choice test.
5. On the final test, within the multiple choice group, the mean test score of high achievers expecting and receiving a multiple choice test will not be significantly different from the mean test score of high achievers expecting a completion test and receiving a multiple choice test.
6. On the final test, within the multiple choice group, the mean test score of low achievers expecting and receiving a multiple choice test will not be significantly different from the mean

test score of low achievers expecting a completion test and receiving a multiple choice test.

Discussion

The greater standard deviations of the completion test results compared to the multiple choice test results for Session 1 and Session 2 suggest that the subjects may have used more guessing strategies to answer the completion questions than to answer the multiple choice questions. If guessing strategies were used to answer the completion questions more often than to answer the multiple choice questions, they did not seem to have a detrimental effect on the reliability of two of the three stories' questions. It was for only one of the stories, that the completion questions were less reliable than the multiple choice questions. It is possible that guessing strategies may have played a role in producing the lower reliability score. Guessing strategies may have been used just as often to answer both the multiple choice and the completion questions, but the chances of the students' guesses to completion questions being judged correct are probably lower than guesses to multiple choice questions. This reasoning seems plausible since multiple choice

questions provide five possible answers and completion questions can usually have many possible answers. Those condition groups that went from a multiple choice test in Sessions 1 and 2 to a completion test in Session 3 showed a decrease in performance; however, this decrease did not exceed the twenty percent decrease that could be accounted for by guessing strategies. Because their differences in performance over the three trials did not extend beyond what could be explained by chance, the null hypotheses concerning those condition groups could not be rejected. It appears that the unexpected switch in test format that some of the subjects experienced in Session 3 was a factor contributing to greater dispersion of scores. Receiving the unexpected test form did lead to greater score dispersion; however, it did not lead to a significant outcome.

Several outcomes of the study support the validity of the researcher's selection criteria of high versus low achievers, as well as the validity of the tests themselves. The consistency of the high achievers outscoring the low achievers in every experimental condition revealed that the test measure did

discriminate between the two achievement groups. The results from the Final Task showed a significant main effect due to condition groups, and a significant main effect due to Achievement when no other significant differences were found. The tests appeared to be successful at delineating performance differences between high and low achievers as the two groups were defined by the researcher.

The prominance of order effects for both the recognition (Session 2) and recall tests (Session 1 & Session 2) indicated the counterbalancing of order was not totally effective. This may be related to using only three (ABC, BCA, or CAB) of the possible six counterbalancing combinations. In the three orders that were chosen, A was always followed by B, B was always followed by C, and C was always followed by A. In part, this may explain the main effects due to order. More variable counterbalancing combinations may have combatted the order effects.

The scores of both achievement groups improved from Session 1 to Session 2 when they received recognition tests. The practice from Session 1 and Session 2 had the same effect on both high and low

achievers. As with the recognition tests, recall practice had an effect on the test scores. Interestingly, the practice from Session 1 to Session 2 resulted in an increase in scores for the high achievers, and a decrease in scores for the low achievers. The main effect due to achievement group suggests that high achievers may have experienced more test expectancy from the practice sessions than the low achievers. The high achievers may have used this expectany to guide their study behavior. Even on the Final Task, the high achievers' results were in the expected direction (Jacoby, 1973) with those subjects who expected and received a completion test outscoring those who expected a multiple choice test and received an unexpected completion test (see Figure 3). Conversely, the main effect due to achievement group also indicates that low achievers may not experience induced expectancy from the practice sessions and therefore lack information to guide their study behavior. On the Final Task, the results of the low achievers were not in the expected direction, with those who expected and received completion tests being outscored by those who expected a multiple choice test

and received an unexpected completion test. Being faced with repeating a similar task produces, perhaps, a lack of both motivation and interest in low achievers. Additionally, the low achievers may have perceived the recall tests as too difficult since a similar decrease in recall performance was not evident for the recognition tests. This interpretation supports Hall et al., (1976) who claimed that differences in performance on recall versus recognition tests may be due either to the amount or to the type of effort perceived by the student. They claimed that the recall task may be perceived as more difficult than the recognition task, thereby requiring more concentrated study effort. The main effects due to achievement found in both the completion and multiple choice data emphasize the consistent differences between high and low achievers in their recall performance. The Final Task results for the recognition data (see Figure 4) were in the expected direction, with those groups who expected and received multiple choice data recalling more than those who unexpectantly received the completion tests. This outcome along with the greater performance difference between the achievement groups

on the completion tests indicates the low achievers may have found the multiple choice tests less difficult than the completion tests.

There were no significant differences found between the subjects expecting and receiving the same test type (Final Task: same) and the subjects expecting one type of test and receiving another (Final Task: different). This also shows a lack of the recall TES effect that has been found in prose studies using cued recall tests (d'Ydewalle, et al., 1982; d'Ydewalle, et al., 1983; Rickards & Friedman, 1978) and recognition tests (May & Piekkola, 1983; Schmidt, 1983). According to Neely and Balota (1981) the number of study-test trials may be important to inducing the test expectancy effect; however, the effect has been achieved with no study-test trial (Schmidt, 1983) and just one study-test trial (Leonard and Whitten, 1983). Past studies also indicated that free study formats, which were used in the present study, may be more likely to produce test expectancy than discrete study formats (d'Ydewalle, et al., 1983).

The Final Task results seem to suggest either that expecting a certain type of test (recall versus

recognition) does not alter the way in which high or low achievers prepare to take that type of test or that any shifts in study by the subjects made no difference to their recall performance. Since the unexpected switch did not result in a significant effect, both groups of achievers may study in the same fashion regardless of test type expected. It may be that both high and low achievers could benefit from instruction on how to use specific study skills for specific test types. Studies teaching categorical and alphabetical organization (Shuell, 1983) and appropriate advance organizers (Tyler et al., 1983) have proven to be especially helpful to low achievers.

Leonard and Whitten's (1983) proposal that recall and recognition processing involved different types of association seems to hold true for both high and low achievers. Recall associations (item-to-item) proved to be more difficult than recognition associations (item-to-context) for both achievement groups. This seems to support Underwood's (1972) and Anderson and Bower's (1972, 1974) suggestion that recognition and recall encoding are qualitatively different.

According to Gillund and Shiffrin's (1984) model,

both recall and recognition performance could be predicted to rise from increased interitem coding. High achievers appeared to demonstrate more interitem coding than did low achievers; however, neither group demonstrated enough interitem coding to reduce self-sampling and produce the recall TES effect.

With regard to Anderson and Bower's associative network and node-tagging theory (1972, 1974), the results of the present study may indicate that high achievers are more able than low achievers to use the qualitative encoding differences necessary to processing for recall versus recognition tasks. For recognition encoding, the high achievers may create more adequate tags between the experimental context and each word as they read and study than low achievers. For recall tasks, low achievers may not create as adequate pathway codes and node-tags as high achievers.

In the review of the literature on high and low achievement, it was suggested that specific differences exist between the two achievement groups. Academic areas of difference included organizational (James & Brown, 1973; Rickards & Hatcher, 1977-78; Shuell, 1983; Tyler et al., 1983), memory (Kail, 1979; Torgesen,

1977), study (Mikulecky & Olshavsky, 1979; Owings et al, 1980; Samuels & Horowitz, 1980; Smiley et al., 1977; Sullivan, 1978) and problem-solving (Kavale & Schreiner, 1979; Larkin et al., 1980) skills. From these studies the author suggested that specific differences may also be found between high and low achievers in a test expectancy situation. The recall performance differences found between the high and low achievers did indicate that the two groups process information for recall and recognition tasks quantitatively differently. The high achievers consistently outscored the low achievers in each test type. The interesting similarity between the two achievement groups was the lack of test expectancy effects for the recall and recognition tasks. Neither achievement group was significantly effected by the unexpected switch in Session 3.

Chapter Five: Conclusions

Major Conclusions

The study revealed that the practice (recall or recognition) from Session 1 and Session 2 was experienced differently by the high and low achievers. On the recognition tests, both high and low achievers improved their recall performance with practice. On the recall tests, the low achievers' recall performance decreased with practice while the high achievers' recall increased with practice. The author suggested that the low achievers' decrease in recall performance may be due to the perceived difficulty of the recall tests (Hall et al., 1976). The major outcome of the study was the lack of the recall TES effect for the recall and recognition tests. Since the high and low achievers were not effected by the unexpected switch in Session 3, it would seem that neither group altered its study patterns to match the test type expected or that any alterations in study patterns made no significant differences.

The recall performance scores on the multiple choice tests were consistently higher than the recall

performance scores on the completion tests for both achievement groups. This may be interpreted as support for the views of Leonard and Whitten (1983), Underwood (1972), and Anderson and Bower (1972,1974) who claimed recall and recognition tasks involved qualitatively different encoding. The high achievers seemed to demonstrate qualitatively and quantitatively more interitem codes (Gillund & Shiffrin, 1984), associative networks and node-tags (Anderson & Bower, 1972, 1974) than did the low achievers. Even though there was a definite difference between the high and low achievers in recall performance on the recall and recognition tests, no difference was shown when it came to the recall TES effect. Brown and Smiley (1978) studied the strategies of children and college students as they attempted to study texts. They found that college students improved their recall of important, but not unimportant, elements of texts with study practice. Similar improvement was found in recall performance of eleventh and twelfth graders but not of fifth through eighth graders. The older students benefited from increased study time because they apparently used it to study the important text segments. The younger

students did not concentrate on the important units because they seemed not know what they were. The subjects of the present study were at the ninth and tenth grade level. They did not demonstrate the recall TES effect, perhaps, because they do not use their study time to prepare differently for recall versus recognition tests. The lack of recall TES effect may be related to the developmental stage of the adolescents, while the difference between high and low achievers in recall performance may be related to the reading comprehension abilities of the two achievement groups.

Limitations

As in any study, this one is not without limitations. The students who participated in this study may not be representative of students elsewhere. The socioeconomic status, ability levels, and ethnicity of the students may not be representative of Grade 9 and 10 high and low achieving students in the same or different school districts.

The selection of the subjects was non-random. Since every member of the population did not have an equal opportunity of being chosen, the generalizability

of the results is limited. Random assignment to treatment groups probably reduced the variance, increasing internal validity, to a certain extent. Moreover, the subjects that were chosen had to meet criteria defined by the researcher and carried out by the school counsellor. Some of the identified high and low achievers may have chosen not to participate in the study. If a signed parental consent form was not returned to the school, the students were not allowed to take part.

One of the assumptions of the study is that the experimenter-constructed questions to the stories were all measuring comprehension. A limitation of using experimenter-constructed tests is the lack of a formal validity measure. In addition, the researcher's operational definition of achievement required the use of student grades. Because the researcher was not given access to this information a later correlation of the test results with student grades was not possible. Fortunately, the tests did seem to discriminate between the two achievement groups as they were operationally defined by the researcher.

The study time of seven minutes per story may have

been too long. May and Piekkola (1983) looked at serial position curves to see how their subjects distributed their study time. They found distinct differences in how their subjects allotted their study time for recall versus recognition tasks when they were given a short amount of time (2.5 minutes). When the subjects were given a longer length of time (5 minutes) there was little difference in how the students distributed their study time for recall versus recognition tests. Seven minutes of study time could have decreased the strength of any TES effects present. Other researchers (May & Piekkola, 1983) also present the test questions in the same order as the information was presented in the study material. Presenting the test questions in an order that was different from how the information appeared in the study material could have increased the overall difficulty of the task, resulting in a further decrease in any TES effects present.

In research projects that involve adolescents as subjects, there can be unique problems. The students were tested in a large group and were asked to try their best; however, peer pressure could have

influenced some adolescents' performance. It was hoped that the presence of a male and female experimenter combatted any experimenter bias due to gender.

A significant limitation in the study was the presence of order effects. It is possible that these effects could have confounded the results, leaving an outcome of no recall TES effect. If there had been more subjects per condition group, additional analyses could have revealed the different effects due to gender and the impact of the order effects in relation to test expectancy. As mentioned earlier, more variable counterbalancing orders could have combatted the order effects more efficiently. An increased number of subjects per cell would have provided a larger N for the reliability (K-R 20) measure, which would give a more powerful indication of the tests' measurement consistency. It is possible that the order effects may have contributed to the large range in reliability coefficients among the tests. Further refinement of the test questions could lead to higher reliability of both the multiple choice and completion tests. This may also lead to less powerful order effects.

Future Research

There are specific recommendations that can be made from the present study. A large source of variance could have come from the testing context itself. In other words, a more controlled testing situation such as a one-to-one or small group test administration could have been more effective. Additionally, subjects from an older age group might have demonstrated the recall TES effect. Since Godard (1985) did report a recall TES when the same stories and questions were used with university students, it would appear that the subjects of the present study were too young to demonstrate the recall TES effect. As mentioned earlier, presenting the test questions in the same order as the information appeared in the study material may have made the task less difficult, resulting in a TES effect. As well, the TES effect may have been magnified with the use of a shorter study time than the seven minutes that was used.

Further study of the TES with adolescents seems warranted. It would be interesting to see if the recall TES effect could be demonstrated in this age group and under what conditions (recall versus

recognition tests). Discovery of the recall TES effect among adolescents could lead to a greater understanding of the developmental factors of test expectancy effects.

Reading ability seems to be an appropriate measurement tool for differentiating high and low achievers. This may be because retention of these materials depends upon reading ability. The investigation of retention when high and low achievers can read the study material equally well could be important to understanding the recall performance differences between the two achievement groups.

It would also be beneficial to find out more about the actual strategies and methods high and low achieving students' use to prepare for different test types. To investigate their study procedures, the use of unobtrusive measures such as serial position effects could be necessary. Serial position effects of high and low achievers could be used to explore the portion of material that each group emphasizes in their study time.

Further study of the achievement variable in relation to test expectancy could reveal important

differences to how each achievement group prepares for different test types. This could lead to an explanation of the qualities that make high achievers different from low achievers. The implications of these findings could result in instructing low achievers in ways to prepare for different types of tests successfully. Because testing concerns millions of people at some point in their lives (Frederiksen, 1984), the continued research into what makes skilled test takers different from less-skilled test takers is a valuable endeavor.

References

Anderson, J.R., & Bower, G.H. (1972). Recognition and retrieval processes in free recall. Psychological Review, 7, 97-123.

Anderson, J.R., & Bower, G.H. (1974). A propositional theory of recognition memory. Memory & Cognition, 2, 406-412.

Ausubel, D.P. (1968). Educational psychology: A cognitive view. New York: Holt.

Balota, D.A., & Neely, J.H. (1980). Test-expectancy and word-frequency effects in recall and recognition. Journal of Experimental Psychology: Human Learning & Memory, 6 (5), 576-587.

Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). Taxonomy of educational objectives. The classification of educational goals. Handbook 1: Cognitive domain. New York: Longmans Green.

Brown, A.L., & Smiley, S.S. (1978). The development of strategies for studying texts. Child Development, 49, 1076-1088.

Connor, J.M. (1977). Effects of organization and expectancy on recall and recognition. Memory & Cognition, 5 (3), 315-318.

Craik, F.I.M., & Lockhart, R.S. (1973). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Duchastel, P.C. (1981). Retention of prose following testing with different types of tests. Contemporary Educational Psychology, 6, 217-226.

d'Ydewalle, G., Degryse, M., & Swerts, A. (1982). Intentional learning in text processing. In A. Flammer & W. Kintsch (Eds.), Discourse processing (pp. 306-313). New York: North-Holland.

d'Ydewalle, G., & Rosselle, H. (1978). Test expectancies in text learning. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), Practical Aspects of Memory. New York: Academic Press.

d'Ydewalle, G. (1981). Test expectancy effects in free recall and recognition. Journal of General Psychology, 105, 173-195.

d'Ydewalle, G., Swerts, A., & De Corte, E. (1983). Study time and test performance as a function of test expectations. Contemporary Educational Psychology, 8, 55-67.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. American Psychologist, 39 (3),193-202.

Fry, E. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. Journal of Reading, 21, 242-252.

Godard, J. (1985). Test expectancy, text organization and question answering in prose text. Unpublished honor's thesis, University of Victoria, Victoria, B.C.

Hall, J.W., Grossman, L., & Elwood, K.D. (1976). Differences in encoding for free recall versus recognition. Memory & Cognition, 4 (5), 507-513.

Hall, J.W., Miskiewicz, R., & Murray, C.K. (1977). Effects of test expectancy (recall vs. recognition) on children's recall and recognition. Bulletin of the Psychonomic Society, 10 (5), 425-428.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1979). Applied statistics for the behavioral sciences. Boston: Houghton Mifflin Co.

Hykin, S. (1983). Recall and recognition test expectancy review. Unpublished manuscript, University of Victoria, Victoria, B.C.

Jacoby, L.L. (1973). Test appropriate strategies in retention of categorized lists. Journal of Verbal Learning and Verbal Behavior, 12, 675-682.

James, T.G., & Brown, B.R. (1973). The effects of prose organization and individual differences on free recall. (Report No. WP-3). Washington, D.C.: Office of Naval Research, Personnel and Training Research Programs Office. (ERIC Document Reproduction Service No. ED 084 873)

Jersild, A.T. (1929). Examination as an aid to learning. Journal of Educational Psychology, 20, 602-609.

Kail, R. (1979). Use of strategies and individual differences in children's memory. Developmental Psychology, 15 (3), 251-255.

Kavale, K., & Schreiner, R. (1979). The reading processes of above average and average readers: A comparison of the use of reasoning strategies in responding to standardized comprehension measures. Reading Research Quarterly, 15 (1), 102-128.

Kintsch, W. (1970). Models for free recall and recognition. In D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.

Leonard, J.M., & Whitten, W.B. (1983). Information stored when expecting recall or recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 440-455.

Marriot, M. (1985). Test expectancy effects and study material selection: Free recall, and recognition of prose text. Unpublished honor's thesis, University of Victoria, Victoria, B.C.

May, R.B., & Sande, G. (1982). Encoding expectancies and word frequency in recall and recognition. American Journal of Psychology, 95, 485-495.

May, R.B. & Piekkola, B. (1983). Test expectancy and question answering in prose processing: A pilot study. Unpublished pilot study. University of Victoria.

Meyer, G. (1936). The effect of recall and recognition on the examination set in classroom situations. Journal of Educational Psychology, 27, 81-99.

Mikulecky, L., & Olshavsky, J.E. (1979). Differences in textbook strategies: Good and poor university level readers perceive themselves employing. Paper presented at the annual meeting of the National Reading Conference, San Antonio, TX.

Naus, M.J., Ornstein, P.A., & Kreshtool, K. (1977). Developmental differences in recall and recognition: The relationship between rehearsal and memory as test expectation changes. Journal of Experimental Child Psychology, 23, 252-265.

Neely, J.H., & Balota, D.A. (1981). Test-expectancy and semantic-organization effects in recall and recognition. Memory & Cognition, 9 (3), 283-300.

Owings, R.A., Petersen, G.S., Bransford, J.D., Morris, C.D., & Stein, B.S. (1980). Spontaneous monitoring and regulation of learning: A comparison of successful and less successful fifth graders. Journal of Educational Psychology, 72 (2), 250-256.

Rickards, J.P., & Friedman, F. (1978). The encoding versus the external storage hypothesis in note taking. Contemporary Educational Psychology, 3, 136-143.

Rickards, J.P., & Hatcher, C.W. (1977-78). Interspersed meaningful learning questions as semantic cues for poor comprehenders. Reading Research Quarterly, 13 (4), 538-553.

Samuels, S.J., & Horowitz, R. (1980, June). Good and poor reader recall of oral and written expository discourse at two levels of difficulty. Paper presented at the annual meeting of the Conference on Reading, Moorhead, MN.

Sax, G., & Collett, L.S. (1968). An empirical comparison of the effects of recall and multiple-choice tests on student achievement. Journal of Educational Measurement, 5 (2), 169-173.

Schmidt, S.R. (1983). The effects of recall and recognition test expectancies on the retention of prose. Memory & Cognition, 11 (2), 172-280.

Shuell, T.J. (1983). The effect of instructions to organize for good and poor learners. Intelligence, 7 (3), 271-286.

Smiley, S.S., Oakley, D.D., Worthen, D., Campione, J.C., & Brown, A.L. (1977). Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. Journal of Educational Psychology, 69, (4), 381-387.

Sowell, E.J., & Casey, R.J. (1982). Research methods in education. Belmont, California: Wadsworth Publishing Co.

Sullivan, J. (1978). Comparing strategies of good and poor comprehenders. Journal of Reading, 21 (8), 710-715.

Torgesen, J.K. (1977). Memorization processes in reading-disabled children. Journal of Educational Psychology, 69 (5), 571-578.

Tyler, S.W., Delaney, H., & Kinnucan, M. (1983). Specifying the nature of reading ability differences and advance organizer effects. Journal of Educational Psychology, 75 (3), 359-373.

Underwood, B.J. (1972). Are we overloading memory? In A.W. Melton & E. Martin (Eds.), Coding processes in human memory. New York: Wiley.

Wrightsman, L.S., & Deaux, K. (1981). Social psychology in the 80's (3rd ed.). California: Brooks/Cole.

Appendix A

Operational Definitions

1. Completion test - a test consisting of fill-in-the-blank questions. It is a form of cued recall (see item 2). The terms completion, recall, and fill-in-the-blank are synonymous in this study.
2. Cued recall - subjects are asked to remember the material with aid, ie., placing words in the context of a sentence (fill-in-the-blank).
3. Discrete study - presentation of one word or item at a time.
4. Free recall - subjects are asked to remember the material in any order without any aids or cues.
5. Free study - presentation of the entire list of words or items to be studied.
6. High achieving student - a student who received a "B" or higher in his English class for at least two out of the three most recent report cards
7. High frequency words - common words, typically those occurring 100 times or more per million in published prose.

8. Interpolated activities - distractor activities placed between the study and test period within a trial. The activities are used to prevent rehearsal of the study material. An example of an interpolated activity would be having the subject count backwards out loud between the study and test tasks.
9. List studies - studies using lists of words (generally nouns) as the study material.
10. Low achieving student - a student who received a "C-" or lower in his English class for at least two out of the three most recent report cards
11. Low frequency words - rare words, often appearing less than 10 times per million words in published prose.
12. Multiple choice test - a test consisting of two or more (generally four or five) alternative answers to choose from. The terms multiple choice and recognition are synonymous in this study.
13. Prose studies - studies using common prose text as the study material.

14. Recall test expectancy - the belief that a completion test will be received, induced by practice trials.
15. Recall test expectancy superiority (TES) effect - people who expect a completion test and get a completion test perform significantly better than people who expect a multiple choice test but receive an unexpected completion test.
16. Recognition test expectancy - the belief that a multiple choice test will be received, induced by practice trials.
17. Recognition test expectancy superiority (TES) effect - people who expect a multiple choice test and receive a multiple choice test perform significantly better than people who expect a completion test but receive an unexpected multiple choice test.
18. Retention score - percent correct; the amount of prose or words remembered.
19. Study-test trial - one study period and one test period equals one study-test trial.

20. Test Expectancy - a phenomenon induced by any or all of the following procedures: a) instruction(s) in the expected test (completion or multiple choice); b) examples of the expected questions; c) one or more practice trials. (From S. Hykin's unpublished review of the test expectancy literature.)
21. Word frequency effect - inferior recall of rare words when tney appear in a list of all rare words.

Appendix B

Parental Consent Form

Study Skills Experiment by Susan Goldsmith

Dear Parent,

I am currently working on my Master's degree in Education at the University of Victoria. Permission for this study has been given by the principal of Cedar Hill and the Superintendent of Greater Victoria School District #61. My study is an experiment on adolescents and how they study for tests. They will be required to read three one-page stories and answer a short fifteen-item quiz on each story. The experiment will take less than one hour in total, and will be done in school on May 28, 1985. The experiment is simple and brief. Yet, it addresses an issue very relevant to your child's education. The study has nothing to do with the students' school requirements and the school will not know their results. If you do not object to your child's participation in the experiment, or to me having access to your child's academic records, please sign this form and have it returned to the school by May 5, 1985. More information about the study can be obtained from me at 386-5972 or 721-7766 before or after the study. Your help is greatly appreciated.

Please return to Mrs. Robinson's drawer in the office.

Parental Consent

Appendix C

Study And Test Materials

LAND DOWN UNDER

Australia is a land far from tne United States. In fact, it is farther from the United States than any other place on earth. Still, it would be hard to find a country in which Americans would feel more at home. Australians seem almost like supertypical Americans. They are gay and robust. They are guileless and good humoured. They bubble with optimism and goodwill. There is an athletic sun bronzed air about them that smacks of Texas. The size and shape of Australia nearly match the contours of the United States turned upside down. The seasons are the reverse of those in the United States. These things add to the overall effect of likeness but with differences. Australians, like Americans, have an innate sense of democracy. Theirs is perhaps even more deeply rooted. It stems from the beginnings shared by the earliest settlers. All were convicts. All were declared freed men. All were to help settle the wild land and develop it for England.

Like North America, the Australian continent is one of wide open spaces. Australia's Wild West was opened up in much the same way as the American West. Cattlemen and sheepmen pushed hard across the Blue Mountains. They headed for the pastures beyond Sydney. There they squatted to stake their claims. There they started their ranches. The soil in New South Wales was fertile. Indeed, it was said that if you tickled tne land with a hoe, it would laugh with a harvest. Then gold was discovered. Within thirty years, the gold rush had tripled the population. The gold rush gave Australia a boomtown glow that it still retains. Australia is an exciting land. It stands alone, surrounded by oceans. Because of this, it provides a home for species of wildlife not known elsewhere. Most of these species have long been extinct in other parts of the world. Perhaps the best known of these are the leaping kangaroos. These are marsupials, animals that carry their young in pouches.

There are also those cuddly teddy bears, the koalas ... And even earlier species like the duck-billed platypus and the anteater abound. The

whole country is an aviary of unusual bird life. There is the kookaburra with its wild, crazy laugh. There is the cassowary, the bird that cannot fly. There are black swans and mutton birds. The emu is second in size only to its cousin, the ostrich. The lyrebird is one of the world's most beautiful warblers. And the miniature penguins, a major attraction at Phillip Island off Melbourne, are a favorite sight of Australians. At twilight they march down to the beach in a fine parade. Among its unique plants Australia boasts the two hundered foot high king karri tree. This is not the only startling tree. Australia also has almost seven hundred types of giant eucalyptus trees. There are wattle trees with golden blossoms. There are bottle trees with trunks shaped like wine casks.

LAND DOWN UNDER

Please answer the following questions.

1. Where did the cattlemen and sheepmen go? _____
2. The soil in New Wales was _____ ?
3. Name the bird that has a wild crazy laugh. _____
4. Which country do the contours of Australia turned upside down look like? _____
5. Name the bird that can not fly. _____
6. Australia's earliest settlers were _____ ?
7. The gold rush gave Australia the boom town _____ it still retains.
8. Animals that carry their young in pouches are called _____ ?
9. The _____ is one of the world's most beautiful warblers.

10. Australians seem almost like _____ Americans.
11. Australians bubble with good will and _____ .
12. What mountains did the cattlemen and sheepmen cross? _____
13. _____ trees have golden blossoms.
14. What tree in Australia grows to be two hundred feet high? _____
15. The settlers were to develop Australia for ____ .

LAND DOWN UNDER

Circle the correct answer (A, B, C, D, E).

1. Where did the cattlemen and sheepmen go?
   A) inland B) pastures C) plateaus D) meadows
   E) lakes

2. The soil in New Wales was _____ .
   A) dark B) rich C) hard D) nutrious
   E) fertile

3. Name the bird that has a wild crazy laugh. _____
   A) cassowary B) kookaburra C) emu
   D) lyre bird E) loon

4. Which country do the contours of Australia turned upside down look like? _____
   A) USA B) Canada C) Greece D) Mexico
   E) Japan

5. Name the bird that can not fly. _____
   A) cassowary B) kookaburra C) emu
   D) lyre bird E) loon

6. Australia's earliest settlers were _____ .
   A) prisoners B) soldiers C) convicts
   D) slaves E) pilgrims

7. The gold rush gave Australia the boom town _____ it still retains.
   A) warmth B) brightness C) excitement
   D) glow E) fervor

8. Animals that carry their young in pouches are called _____ .
   A) mammals B) wallabys C) kangaroos
   D) platypuses E) marsupials

9. The _____ is one of the world's most beautiful warblers.
   A) cassowary B) kookaburra C) emu
   D) lyre bird E) loon

10. Australians seem almost like _____ Americans.
A) supersimilar B) archetypical C) typical
D) pseudo E) supertypical

11. Australians bubble with good will and _____ .
A) wonder B) optimism C) hope
D) laughter E) humour

12. What mountains did the cattlemen and sheepmen cross? _____
A) gold B) white C) green D) blue
E) black

13. _____ trees have golden blossoms.
A) wattle B) bottle C) walnut D) oak
E) eucalyptus

14. What tree in Australia grows to be two hundred feet high?
A) king korri B) birch C) oak D) king karri
E) red cedar

15. The settlers were to develop Australia for _____ .
A) farming B) Europe C) England D) ranching
E) cultivation

LITTLE PLANE, BIG ATLANTIC

The day came listlessly, limp with rain, as the crowd waited in the mud of Roosevelt Field on Long Island. Then the silver-colored plane was wheeled onto the runway. The rain had ceased. Shortly after seven o'clock, that morning of May 20, 1927, the Spirit of St. Louis seemed ready to go. As Charles Lindbergh, the pilot, stepped from a closed car, the crowd of several thousand stirred expectantly. They struggled to get a look at the youthful pilot. Lindbergh was twenty-five, but in the eyes of the crowd he appeared a boy. He looked at the plane. The wheels pressed deeply into the soft wet clay of the runway. The plane bore a heavier fuel load than the Spirit of St. Louis had ever lifted. Above a hanger a wind sock puffed. The wind was shifting from east to west. He would soon have to contend with a tailwind which would increase the dangers of the take-off. He climbed into the cockpit.

A mechanic said, "She's thirty revolutions low." Lindbergh let the blades idle and sank back into his seat. Wind, weather, power, load - they all appeared to be working against him. The crowd waited. Few would blame him if he decided to leave on another day. Then the take-off would be less risky. But wasn't this his chance to beat his rivals in the race across the Atlantic? The race had really started in 1919. A New York hotel owner then said that he would give $25,000 for the first non-stop flight between New York and Paris. For seven years the 3600 mile flight seemed an impossible feat. By 1926, however, plane engines and designs had improved a great deal. The race was on. The pursuit of the prize had already cost six lives. Four men had died when their planes had crashed during take-off. Two more had flown out over the ocean and were never seen again. A navel search failed to locate them.

And now a new attempt for the prize was being made. But this time it was being sought by only one man. The young man seated alone in the Spirit of St. Louis had the runway to himself. At 7:52 A.M. Lindbergh buckled his safety belt. The Spirit of St. Louis moved slowly at first, then gathered speed.

Ahead, a white cloth on a stick marked the point where Lindbergh would have to take off. Beyond it he couldn't stop without crashing. Tense and still, the crowd waited. The Spirit of St. Louis sped on. The wheels rose from the ground, then settled again. The crowd groaned. In the cockpit, the pilot pulled the stick back firmly and the plane rose once more into the air. It flew over a tractor at the end of the runway. It cleared some telephone wires by twenty feet, and swept along over the houses. Then it climbed over a hilltop and into a curtain of mist. Beyond the curtain lay Paris.

LITTLE PLANE BIG ATLANTIC

Please answer the following questions.

1. The plane cleared the telephone wires by _____ feet.
2. How many men died during take off? _____
3. A mechanic said that the engine was _____ revolutions low.
4. The plane climbed over a hilltop and into a _____of mist.
5. The silver colored plane was _____ on to the runway.
6. The dangers of takeoff were increased by a _____wind.
7. The pilot pulled back _____ on the stick.
8. Ready to fly, at _____ Lindbergh buckled his safety belt.
9. By _____ , plane engines and designs had improved a great deal.

10. The crowd waited in _____ field on Long Island.
11. In what year was Lindbergh ready to make his flight? _____
12. Wind, weather, power, and _____ were working against Charles.
13. The plane flew over a _____ at the end of the runway.
14. When had the race to cross the Atlantic started? _____
15. Above a hanger a wind sock _____ .

LITTLE PLANE BIG ATLANTIC

Circle the correct answer (A, B, C, D, E).

1. The plane cleared the telephone wires by _____ feet.
   A) 12 B) 14 C) 16 D) 18 E) 20

2. How many men died during take off? _____
   A) 1 B) 2 C) 3 D) 4 E) 5

3. A mechanic said that the engine was _____ revolutions low.
   A) 20 B) 25 C) 30 D) 35 E) 40

4. The plane climbed over a hilltop and into a _____ of mist.
   A) bank B) screen C) veil D) curtain E) shadow

5. The silver colored plane was _____ on to the runway.
   A) wheeled B) taxied C) flown D) rolled E) pushed

6. The dangers of takeoff were increased by a _____ wind.
   A) westerly B) tail C) easterly D) strong E) steady

7. The pilot pulled back _____ on the stick.
   A) heavily B) rigidly C) firmly D) tensely E) frightfully

8. Ready to fly, at _____ Lindbergh buckled his safety belt.
   A) 7:35 B) 7:49 C) 7:52 D) 7:57 E) 8:04

9. By _____ , plane engines and designs had improved a great deal.
   A) 1936 B) 1924 C) 1926 D) 1934 E) 1933

10. The crowd waited in _____ field on Long Island.
A) Flanders B) Johnson C) Roosevelt
D) Richardson E) Eisenhower

11. In what year was Lindbergh ready to make his flight? _____
A) 1920 B) 1923 C) 1925 D) 1927 E) 1933

12. Wind, weather, power, and _____ were working against Charles.
A) time B) energy C) distance D) force
E) load

13. The plane flew over a _____ at the end of the runway.
A) barn B) tree C) bailer D) truck
E) tractor

14. When had the race to cross the Atlantic started? _____
A) 1919 B) 1916 C) 1914 D) 1921 E) 1924

15. Above a hanger a wind sock _____ .
A) filled B) puffed C) stirred D) flew
E) inflated

THE MAGIC OF LANGUAGE

During the Third Century ater the birth of Christ, a strange sort of writing appeared. It was in the area that is now southern Denmark and northern Germany. At first glance, the writing didn't resemble anything. There were just angular marks, such as a child might make in the mud with a pointed stick. Certainly it wasn't a bit like the dignified Roman alphabet that was destined to replace it in time. But it was an alphabet, and these early people used it at the times when they wrote. The characters of this alphabet runes must indeed have appeared magical to these early people. But when radio and television seemed magical and strange to us when they were first introduced. These people used the runes not only when they wrote but also when they performed weird, magical rites. The runes were seen not just as an aid to communication. They were magical signs with real power of their own.

There were good runes that brought good luck, bitter runes that wrought evil, protective runes that safeguarded, medical runes that healed. Because runes were linked with the magical rites of heathen worship, their use was forbidden by the Christians and their popularity declined. Many objects of the past have been found that were marked with runes. Of all the ancient objects that bore runic writing, none fascinate the scholar more than the Golden Horns. Their tale begins in Denmark, August 2, 1639. On that day a little girl named Kirsten Svensdatter was walking on a road near the small town of Gallehus. She tripped over what she thought was a tree root. The "tree root" proved to be no such thing. It was found instead to be a dirty, metal object. Washed in a nearby brook, it was seen to be made of bright yellow metal and to show strange figures at play in a pagan festival.

On the advice of friends, Kirsten showed the object to the town goldsmith. He told her that it was a horn of pure gold. Town officials advised Kirsten that she could name her reward. She requested, and received, a red silk skirt. The King gave the horn to his son, who was a collector of antique objects. The Prince used the horn to drink toasts on important occasions. When he died a few years later, the horn

was placed in the Royal museum, in Copenhagen. Many questions concerning the horns remain unanswered. What were they used for? What did they mean? The runes were not used long. Christianity helped bring about their early end. One more cause was the evolutionary development of the Latin alphabet. In general, as civilization progressed, runic writing was quite properly replaced by the more standardized, dignified, useful Roman alphabet. Yet, with the passage of the runes, much of the glamor, mystery, and magic that we associate with writing passed away forever.

THE MAGIC OF LANGUAGE

Please answer the following questions.

1. Kirsten received a red silk _____ as a reward.
2. The use of runes was forbidden by _____ ?
3. In what century after the birth of Christ did a strange sort of writing appear? _____
4. The tree root was in fact a dirty metal object which the girl washed in a nearby _____ .
5. What kind of festival was depicted by the figures on the object?_____.
6. The tale of the Golden Horns began in _____ .
7. The evolutionary development of the _____ alphabet helped bring about the end of the use of runes.
8. Who did Kirsten give the magical object to? _____
9. What does the word "runes" mean? _____

10. The runes were used for a _____ time.
11. Kirsten _____ tripped over a "tree root".
12. The horns were kept in the _____ museum.
13. This writing appeared in the regions that are now Northern Germany and _____ Denmark.
14. Runes were not only magical signs but were also an aid to _____ .
15. Runes were used for writing and performing magical _____ .

THE MAGIC OF LANGUAGE

Circle the correct answer (A, B, C, D, E).

1. Kirsten received a red silk _____ as a reward.
   A) handerkerchief B) blouse C) ribbon D) dress
   E) skirt

2. The use of runes was forbidden by _____ ?
   A) scholqrs B) heathens C) bishoβs D) Romans
   E) Christians

3. In what century after the birth of Christ did a strange sort of writing appear? _____
   A) first B) second C) third D) fourth E) fifth

4. The tree root was in fact a dirty metal object which the girl washed in a nearby _____ .
   A) stream B) creek C) well D) brook E) lake

5. What kind of festival was depicted by the figures on the object? _____
   A) strange B) fertility C) pagan D) merry
   E) religious

6. The tale of the Golden Horns began in _____ .
   A) 1639 B) 1759 C) 1749 D) 1629 E) 1739

7. The evolutionary development of the _____ alphabet helped bring about the end of the use of runes.
   A) Latin B) Roman C) Greek D) Phoenician
   E) Egyption

8. Who did Kirsten give the magical object to? _____
   A) prince B) king C) queen D) princess
   E) magician

9. What does the word "runes" mean? _____
   A) ruins B) letters C) mysteries D) characters
   E) words

10. The runes were used for a _____ time.
A) very short B) short C) lengthy D) long
E) very long

11. Kirsten _____ tripped over a "tree root".
A) Swendatter B) Svenson C) Svendatten
D) Sveenden E) Svensdatter

12. The horns were kept in the _____ museum.
A) Copenhagen B) Prague C) Regal D) Royal
E) State

13. This writing appeared in the regions that are now Northern Germany and _____ Denmark.
A) Western B) Southern C) Eastern D) Central
E) Northern

14. Runes were not only magical signs but were also an aid to _____ .
A) the Romans B) the Pagans C) communication
D) writing E) language

15. Runes were used for writing and performing magical _____ .
A) sacraments B) acts C) rites D) ceremonies
E) dances

Appendix D

Procedure Instructions and Debriefing Details

Experimental Room (Cafeteria):

The following outline was written on the chalkboard:

1. Are you sure of your ID number?
2. Do you have a pen or pencil?
3. You will read 3 stories and answer 3 fifteen item tests altogether.
4. You have 7 minutes to study each story.
5. You have 5 minutes to answer each test.
6. This is repeated 3 times.
7. Only your pen(cil), and brown envelopes with stories and tests should be on your desk.
8. Never take out or turn pages until you are requested to do so.
9. You are allowed to make pen (cil) marks on the study papers.
10. Take top two pages from envelope to your left and put it away in envelope to your right.

Verbal Instructions:

1. Introduction

Good afternoon! I am Susan Goldsmith and this is Michael Marriot. We are working on a project based at the University of Victoria. The project is concerned with how people study for tests - that is, how they study printed text material and how they answer questions about what they have read.

First of all: Are you sure of your ID number? Do you have a pen or pencil?

The study includes three one-page stories and three fifteen-item tests to each story.

Second: You will read three stories and answer three fifteen item tests altogether. You have seven minutes to study each story. You have five minutes to answer each test. This is repeated three times.

The results of this experiments will not effect your school grades. The results of the experiment depend on your performance so it is essential that you try your best and concentrate on what you are doing.

2. Instructions:

When I tell you (not now), you will reach into the brown envelope on the left side of your desk. You will take out the top two sheets and place them in front of you. The cover sheet will say "Story 1 - Do Not Turn Page Until Requested". This is the only thing that should be on your desk besides the envelopes and your pen or pencil.

Please note that the stories and questions have been assigned randomly so it is very unlikely that the person sitting next to you has the same story or test.

Third: Only your pen or pencil, and brown envelopes with stories and tests should be on your desk. Never take out or turn pages until you are requested to do so.

Remember, you will be given seven minutes to study the story and then you will be asked to put it in the envelope on the right side of your desk.

Review: You have seven minutes to study each story and five minutes to answer each test.

You are encouraged to study the way you would normally study for a test. You are allowed to mark

the story pages (point nine of the outline).

We will tell you when you have two minutes and one minute left in the study and test periods. If you are finished early, please sit quietly and wait until the time is up.

Are there any questions??

Remember, try your best!!

3. Experiment:

1a) Take out the top two pages from the envelope on the left side of your desk. The cover should say "Story 1 - Do Not Turn Page Until Requested". Everyone ready? Begin!

2 minutes left.....

1 minute left.....

STOP!

Put the story in the envelope on the right side of your desk.

b) Take out the top two pages from the envelope on the left side of your desk. The cover should say "Test 1 - Do Not Turn Page Until Requested". Remember you have 5 minutes to answer the questions. Even if you are not sure of the answer, put down your guess. <u>Try your best!</u>

Everyone ready? Begin!

2 minutes left.....

1 minute left.....

STOP!

Put the test in the envelope on the right side of your desk.

2 a) & b) Repeat above.

3 a) & b) Repeat above.

Circulated around room during study and test periods.

4. Debriefing:

The following debriefing talk was given with the aid of a transparency and overhead projector. The transparency resembled the top half of Figure 1 of this thesis. Distinctions between high and low achievers were not shown. The students were told the following:

As you know there were three sessions, each session consisting of one story and one test. All of you were divided into two groups. One group received multiple choice questions to the first two stories, and the other group received completion questions to the first two stories. In Session 3

half of the people who received multiple choice tests in the first two sessions were surprised with completion questions to the third story. Similarly, half of the people who received completion questions to the first two stories were surprised with multiple choice questions to the third story. Those people who experienced the same type of test for all three stories were in the control group and those people who experienced the unexpected switch in session three were in the experimental group. After you had received the same type of test twice, we assumed you were expecting the same type of test to story three. The main point of the study was to find out if expecting a certain type of test guided your preparation for that test. We wanted to see if receiving the type of test you expected, resulted in better performance than receiving a different type of test from what you expected. Does anyone have any questions? (The comparison of high and low achievers was not discussed. The Committee on Human Subjects agreed with the researcher that the students did not need to know about this analysis

to understand the general purpose of the study.)

5. Conclusion:

Thank-you very much for participating! Please remain here until the bell goes and then go to your next class.

VITA

Surname: Goldsmith Given Names: Susan Marie

Place of Birth: Chilliwack, B.C

Date of Birth: December 10, 1961

Educational Institutions Attended, with Dates of Entering and Leaving:

OKANAGAN COLLEGE, KELOWNA, B.C. 1979 to 1980

UNIVERSITY OF VICTORIA, VICTORIA, B.C. 1980 to 1983

UNIVERSITY OF VICTORIA, VICTORIA, B.C. 1984 to 1985

____ to ____

Degrees, Diplomas, Etc., Awarded, with Dates and Names of Institutions:

B.A. (Major) 1983 University of Victoria, B.C.

Honors and Awards:

Cominco Scholarship, 1979

University of Victoria Supplement, 1984

E.R.I.B.C. Grant, 1985

Publications:

PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my thesis or dissertation (the title of which is shown below) to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, of similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis/Dissertation

Test Expectancy (Recall and Recognition): It's Effect On Low And High Achieving Grade 9 And 10 Adolescents

Author [redacted]
(Signature)

SUSAN M. GOLDSMITH

October 25, 1985.
(Date)