

Fall 1986

WHAT IS IT THAT YOU KNOW WHEN  
"YOU KNOW"? COMPREHENSION GOALS  
AND COMPREHENSION MISMATCHES IN  
COLLEGE STUDENTS (STUDY BEHAVIORS,  
METACOMPREHENSION, TEST TAKING)

MARCY LOIS KRAUS

*University of New Hampshire, Durham*

Follow this and additional works at: <https://scholars.unh.edu/dissertation>

---

**Recommended Citation**

KRAUS, MARCY LOIS, "WHAT IS IT THAT YOU KNOW WHEN "YOU KNOW"? COMPREHENSION GOALS AND COMPREHENSION MISMATCHES IN COLLEGE STUDENTS (STUDY BEHAVIORS, METACOMPREHENSION, TEST TAKING)" (1986). *Doctoral Dissertations*. 1488.

<https://scholars.unh.edu/dissertation/1488>

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

## INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.



8701338

**Kraus, Marcy Lois**

WHAT IS IT THAT YOU KNOW WHEN "YOU KNOW"? COMPREHENSION  
GOALS AND COMPREHENSION MISMATCHES IN COLLEGE STUDENTS

*University of New Hampshire*

Ph.D. 1986

University  
Microfilms  
International 300 N. Zeeb Road, Ann Arbor, MI 48106



WHAT IS IT THAT YOU KNOW  
WHEN "YOU KNOW"?  
COMPREHENSION GOALS AND COMPREHENSION MISMATCHES  
IN COLLEGE STUDENTS

BY

MARCY L. KRAUS

B.S. (Psychology), Viterbo College, 1979  
M.A. (Psychology), University of New Hampshire, 1982

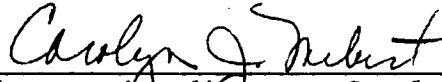
A DISSERTATION

Submitted to the University of New Hampshire  
In Partial Fulfillment of  
the Requirements for the Degree of

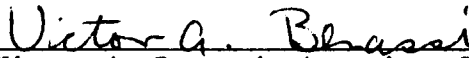
Doctor of Philosophy  
in  
Psychology

September, 1986

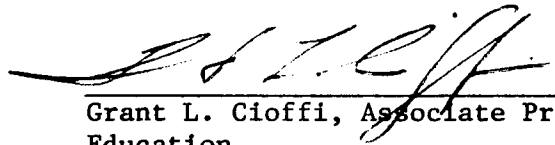
This dissertation has been examined and approved.



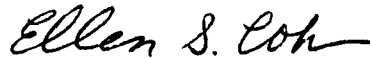
\_\_\_\_\_  
Dissertation director, Carolyn J. Mebert,  
Associate Professor of Psychology



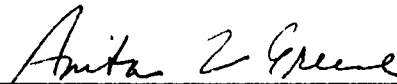
\_\_\_\_\_  
Victor A. Benassi, Associate Professor of  
Psychology



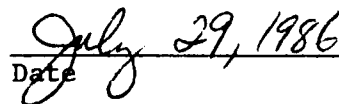
\_\_\_\_\_  
Grant L. Cioffi, Associate Professor of  
Education



\_\_\_\_\_  
Ellen S. Cohn, Associate Professor of  
Psychology



\_\_\_\_\_  
Anita L. Greene, Assistant Professor of  
Psychology

  
\_\_\_\_\_  
Date

## ACKNOWLEDGMENTS

I wish to express my appreciation to those who have been there with me on this project. I would like to thank Carolyn Mebert, my chairperson, for encouraging me to think and grow. I am grateful for her insightful comments and critical analyses--not to mention her patience--as this process has finally reached its conclusion. Ellen Cohn has provided extensive statistical and editorial advice. I am appreciative of her unlimited availability and her friendship; her contribution to this project has been invaluable. I would like to thank Grant Cioffi for his assistance in helping a novice tackle the reading and studying literature. I am also grateful to Victor Benassi for his practical and theoretical observations, and to Anita Greene for her editorial comments.

Many friends have made this project easier to endure. I would like to especially thank Jim Grosch, Jerry Jennings, Maria McKenna, Susan Newman, Judi Scharff, and Deb Kraus for their continual support and encouragement. Jim has helped me reach my final deadline in the completion of this manuscript and Deb, true to her profession, has been a good counselor.

I would like to thank the approximately 400 psychology students who have participated in my research, not to mention those students who, without knowing it, provided the impetus for this project by talking to me about their test-taking and studying difficulties.



I would like to thank the University of New Hampshire and the Department of Psychology for the assistantships and summer fellowships I have received. Without this support, I could not have brought this project to completion. I am also very grateful to my colleagues at the Whittemore School for recognizing my need to disappear into my office occasionally, and for allowing me to hoard the computer terminal.

My family has long remained interested in and supportive of my work. I doubt that I would have conceived of this particular path had I not grown up with an important role model, my mother. Finally, I would like to thank my husband Ray who has never known the relative tranquility of living with a spouse who is not a graduate student. He has never once asked me when I would finish and has always known that I would.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	iii
LIST OF TABLES.....	vi
ABSTRACT.....	vii
CHAPTER	PAGE
I. INTRODUCTION.....	1
Introduction to the Problem.....	1
Metacognition and Comprehension Monitoring.....	3
Defining Understanding.....	6
II. REVIEW OF THE LITERATURE.....	17
Models of Learning.....	17
Subject Characteristics.....	20
Orienting Tasks.....	21
Criterial Tasks.....	25
Characteristics of the Material.....	37
Overview and Statement of Purpose.....	39
III. STUDY ONE.....	44
IV. STUDY TWO.....	56
V. DISCUSSION.....	79
Comprehension Goals.....	79
Quiz Condition Differences.....	84
Dualism and Relativism.....	85
Mismatch Experiences.....	87
Post-Quiz Evaluations and Mismatching.....	89
Reader Interest.....	90
Directions for Future Research.....	91
LIST OF REFERENCES.....	93
APPENDICES.....	119
APPENDIX A-COMPREHENSION GOAL MEASURE.....	120
APPENDIX B-ADHERENCE MEASURE.....	122
APPENDIX C-SIGNIFICANCE TABLES.....	123

## LIST OF TABLES

1. Means and Standard Deviations for Measures by Gender.....	102
2. Percentage of Responses for each Comprehension Goal Level by Response Category.....	103
3. Correlations Between Ratings and Mismatch Frequency.....	104
4. Means and Standard Deviations for Performance Measures by Mismatch Groups.....	105
5. Intercorrelations Among Ratings and Study Behaviors.....	106
6. Reasons for Mismatches.....	107
7. Intercorrelations Among Ratings, Reading, Mismatch and Comprehension Goals.....	108
8. Means and Standard Deviations for Performance Measures by Gender.....	109
9. Factor Analyses of Comprehension Goal Items.....	110
10. Correlations Between Factual and Main Idea Goals, NDRT, and Quiz Scores.....	111
11. Means and Standard Deviations for Performance Measures by Mismatch Group.....	112
12. Correlations Between Quiz Scores and Performance Measures by Mismatch Group.....	113
13. Stepwise Multiple Regression Analyses for the Prediction of Quiz Performance by Mismatch Group.....	114
14. Stepwise Multiple Regression Analyses for the Prediction of Quiz Performance by Quiz Type.....	115
15. Means and Standard Deviations for Post-Quiz Measures by Quiz Condition.....	116
16. Stepwise Multiple Regression Analyses for the Prediction of Post-Quiz Evaluations.....	117
17. Post-Quiz Evaluations.....	118

## ABSTRACT

### WHAT IS IT THAT YOU KNOW WHEN "YOU KNOW"? COMPREHENSION GOALS AND COMPREHENSION MISMATCHES IN COLLEGE STUDENTS

by

Marcy L. Kraus  
University of New Hampshire, September, 1986

The purpose of the present research was to investigate the comprehension mismatch experience as it related to students' comprehension goals and testing experiences. In Study 1, subjects reported their reading-related comprehension goals while studying a 3500-word text. A questionnaire was administered assessing comprehension of and interest in the text, frequency of mismatches and general studying behavior. Results of Study 1 indicated that eighty-seven percent of the students surveyed reported moderate to frequent mismatches. Subjects varying in mismatch frequency did not differ significantly with regard to any of the performance measures. The comprehension level quiz items were rated as most appropriate and the application-analysis items as least appropriate using the criteria established by Bloom et al. (1956).

In Study 2, subjects completed a quiz containing either factual or applied questions after reading a text and reporting their comprehension goals. The results from Study 2 indicated that comprehension goals were not significantly related to overall quiz performance. Stepwise multiple regression analyses yielded different sets of predictor

variables for each quiz type. Subjects' ratings of their mismatch episodes were not significantly related to text comprehension, reading proficiency or quiz performance. It was observed, however, that low and moderate mismatch subjects were more accurate in estimating their test readiness than high mismatch subjects. Post-quiz evaluations indicated that subjects receiving the applied quiz reported significantly higher mismatch ratings of the testing experience than subjects receiving the factual quiz.

The present findings point to the importance of reader interest, reading proficiency, reader expectations and the type of comprehension measure administered, in examining questions related to reading and testing. Continued investigation of the comprehension mismatch phenomenon appears warranted given the high proportion of students reporting such experiences. Future research on comprehension mismatches might be best directed toward developing profiles of subjects differing in mismatch experiences with particular attention paid to learning styles, testing situations related to mismatching, and previous educational experience.

## CHAPTER 1

### INTRODUCTION

#### Introduction to the Problem

"A very basic form of self-awareness is the realization that there is a problem of knowing when you know and when you do not" (Brown, 1980, p.458). Intuitively it would seem that awareness of what it is you know and do not know is a very basic precursor of more complex forms of thought. That there is difficulty in this type of introspection, however, can be predicted on the basis of research from education and psychology. Not only does comprehension remain a difficult term to define (Markman, 1981; Ornell, 1979), learners may differ in their orientations toward and interpretation of "understanding" (Perry, 1970; Ryan, 1984; Svensson, 1977) and instructors may not make explicit, and/or actively demonstrate, how to attain various kinds of comprehension (Anderson & Armbruster, 1984; Kunen, Cohen & Solman, 1981).

In many ordinary circumstances people may be able to sufficiently judge the adequacy of their comprehension. In academic settings, however, comprehension criteria may be especially variable and frequently vague. All that is certain is that comprehension is associated with the requirement to perform particular criterion tasks. If these tasks are not clearly specified, or if subjects lack the

knowledge, skills, or experience to learn information in a particular way, performance on these tasks may be unsuccessful. Of interest to the present research is the role that subjective and individual comprehension criteria play in text processing. The focus is on the question of students' knowledge and expectations regarding their own comprehension activities. Specifically, the following questions will be addressed:

1. When students read material for the purpose of understanding, how do they define "understanding?" That is, given a reading task, what types of comprehension criteria are used by students? How are these criteria related to individual differences in reading ability, orientations toward knowledge (e.g., Perry, 1970), and expectations concerning the nature of the criterion task?

2. How often are there discrepancies between a student's self-reported feelings of comprehension and actual performance on some measure of comprehension? These discrepancies are defined in the present research as "comprehension mismatches." The comprehension mismatch experience has received some attention in the literature but has not been researched extensively (Bransford, 1979; Rubin, 1985). Several issues are considered: (a) how frequently do mismatches occur in the general experience (i.e., perception) of college students; (b) how have these mismatches been interpreted; and (c) what is the relationship between a student's reported mismatch experience and his or her comprehension strategies, reading ability, and other related text-processing variables?

### Metacognition and Comprehension Monitoring

A very general but nevertheless practical definition of studying was provided in 1926 by Butterweck who defined studying as a "pupil activity of the type required to satisfy the philosophy of education held by the teacher" (Butterweck, 1926). As it pertains to the research in reading comprehension, studying has been described as a special form of reading "[involving] all the activities of reading for meaning and more" (Baker & Brown, 1984b, p. 367). These additional activities are those strategies necessary to make sure that the material is not only comprehensible, but relevant and memorable. This includes, for example, determining the purpose of the task, selecting effective study strategies, identifying the important aspects of the information in relation to the criterion task, and recognizing when the material has been mastered. These activities include both cognitive and metacognitive skills. As Flavell (1981) suggested, much of what we think we know involves "at least some monitoring of cognitive goals, experiences and actions" (Flavell, 1981, p. 39). It is this monitoring, specifically identified as cognitive monitoring, which is responsible for the planning, checking, testing and evaluating strategies necessary for achieving understanding.

Recently the term "metacognition" and its offspring metacomprehension, comprehension monitoring, and study monitoring, have been used to describe the particular types of knowledge one may have about some cognitive endeavor. A hierarchical relationship between these terms has been postulated such that metacognition can be viewed as superordinate to the more specific terms, comprehension monitoring and



study monitoring. While metacognition refers to knowledge about cognitions in general, comprehension and study monitoring are viewed as applying more specifically to reading comprehension (Baker & Brown, 1984a).

The question posed by the title of this dissertation, "What is it that you know when 'you know'?" is essentially a metacognitive question. It, and its companion question, "When do you know when 'you know'?" would appear to mediate most successful studying efforts. Such monitoring activity is expected to lead to a wide variety of experiences concerning tasks, goals and strategy use. While monitoring progress toward a goal, questions such as these may prompt the learner to try actively to figure out how to remedy comprehension problems, to assess readiness for a test or to revise an original study plan.

Following the introduction of the term metacognition in 1970 (Flavell, Friedrichs, & Hoyt, 1970), Flavell proposed a model of cognitive monitoring composed of four basic components--cognitive goals, cognitive actions, metacognitive knowledge and metacognitive experiences. The components themselves include knowledge about oneself as a learner (metacognitive knowledge), feelings and cognitions pertaining to the progress of the cognitive endeavor (metacognitive experiences), the goal or purpose of the cognitive enterprise (cognitive goals), and the strategies undertaken to achieve these goals (cognitive actions). Each component is believed to exert influence on or receive input from other components.

For example, once a learner has identified the particular goals of a task (e.g., prepare for a test in biology), the learner's metacognitive knowledge, or what Flavell (1981) has described as that part of "your accumulated world knowledge that has to do with people as cognitive agents and with their cognitive tasks, goals, actions, and experiences" (p. 40), directs the learner in identifying how the cognitive task should best be undertaken. Particular cognitive actions (i.e., study strategies) are selected in accordance with the learner's existing knowledge about her own learning abilities and expectations of the task demands. When progress or lack of progress toward the goal is consciously recognized, it is in the form of a metacognitive experience. Metacognitive experiences are ideas, thoughts and affects related to some aspect of the cognitive enterprise:

We feel that we have just begun, or that the goal is still very remote. We feel we are almost there...We feel that we are in fact there...Or we feel that the goal is possibly, probably, or definitely unattainable" (Flavell, 1981, p. 48).

In turn, these metacognitive experiences are expected to have important effects on both cognitive goals and actions.

The general definition of metacognition as knowledge about one's cognitions (Brown, 1978), has been more specifically depicted as involving two separate (not necessarily independent) activities: (a) an awareness of the skills, strategies, and resources needed to perform a task effectively; and (b) the ability to control and monitor these mechanisms successfully (Baker & Brown, 1984a). Since successful readers must have some awareness and regulation of the processes they engage in as they read, theories describing the reading process

typically include a discussion of the skills and activities that involve metacognition (Baker & Brown, 1984b). The current popularity of all the new "meta" words, however, is not to imply that such a focus is anything new or revolutionary. Early conceptualizations of reading also point to the importance of the planning, regulating, and evaluating activities currently referred to as metacomprehension (e.g., Dewey, 1910; Thorndike, 1917). What appears to be new, however, is an interest in metacognition per se as it applies to a number of cognitive realms such as communication, problem solving, reading, and studying. As Brown (1980) suggests:

One major justification for studying metacognitive skills is that they do appear to have 'ecological validity'; that is, there are recognizable counterparts in 'real-world, everyday life' situations. Checking the results of an operation against certain criteria...is a metacognitive skill applicable whether the task under consideration is solving a math problem, reading for meaning, memorizing a prose passage, following a recipe, or assembling an automobile or piece of furniture (p. 454).

#### Defining Understanding

Knowing when you do or do not understand as applied to reading and studying would appear to be a vital component of academic achievement. Making this judgement, however, may not always be particularly straightforward. Studying involves reading in preparation for some criterion task. The kind of task may be known in varying degrees, but unless students have a duplicate copy of the measure to be administered, there is typically some element of uncertainty. As every instructor knows, when students say "yes, I read that; yes, I understand that," it is unclear what that "yes" really means. To some students it may mean "memorize the definitions," to other students it may mean "get the gist"

and to still others it may mean "I know it because I spent four hours reading it." Because the criteria for knowing when one has achieved comprehension are sometimes vague, individuals may have difficulty applying them; in addition, when they do apply them, such definitions may vary according to individuals, tasks, and settings.

Starting with the basic question "what is comprehension?" the following insights by students, theorists, and researchers, provide a brief, by no means definitive, spectrum of responses. Those whose goal it is to comprehend material in order to pass undergraduate courses have said the following in response to the question "How do you know when you adequately comprehend material for a course or test?" "If I am able to explain it in my own words." "If the text has questions in it I see how well and quickly I can do the problems." "When I can memorize all my notes without looking." "If I can apply the material to other courses...I know I haven't just memorized it." "When I'm completely saturated and cannot (will not) absorb anymore." "[When I can] go through a complete rundown of the material as if I were giving a lecture." "When I can explain the information to someone else comfortably..." "I feel confident and am not worried." "When I read something I either understand it or I don't, and whichever one it is, I know it (it's a gut feeling)." "I don't. I won't know until I take the test." (Kraus, 1984).

As demonstrated by the previous responses, for students, knowing if one understands is often determined with reference to particular activities believed to be accurate indicators of this phenomenon. Comprehension "by default" would seem to be obtained when there are no

obvious signs to the contrary. Feelings are probably important but not always sufficient; as one student remarked, "I get a feeling I know what I'm talking about, and proceed to do terribly on the test. Other times I feel terribly anxious before a test (feel like I don't know it) and do well." Such mismatches appear to be common among undergraduates (Kraus, 1984). As Baker and Brown (1984b) have pointed out

Readers who understand incorrectly have much the same feelings as readers who understand correctly. Hence, they can hardly be expected to take remedial action when comprehension fails, since they don't realize that comprehension has in fact failed (p. 356).

Most reading theorists would agree that comprehension can be characterized as an active, constructive process (Baker & Brown, 1984b). For example, in an article titled "Reading as reasoning," Thorndike (1917) suggested that

Understanding a paragraph is like solving a problem in mathematics. It consists in selecting the right elements of the situation and putting them together in the right relations. The mind is assailed as it were by every word...It must select, repress, soften, emphasize, correlate and organize, all under the influence of the right mental set of purpose or demand" (p. 329).

Comprehension also typically involves more than recognizing a string of words and sentences: "Learning from a text does not mean rote memory...Learning from reading often requires thinking and making inferences if true comprehension is the goal" (Gibson & Levin, 1975, pp. 435-436). The ability to draw inferences, to evaluate, to solve problems, is likewise related to the reader's current knowledge and experiences. The importance of existing knowledge structures has been identified as an integral component of comprehension:

Whether we are aware of it or not, it is [the] interaction of new information with old information that we mean when we use the term comprehension. To say that one has comprehended a text is to say that [one] has found a mental 'home' for the information in the text" (Anderson & Pearson, 1984, p. 255).

This notion of "meaning in the person" as important in cognitive activities has long been recognized by philosophers and psychologists. E.B. Titchener was perhaps one of the earliest psychologists to come to scientific terms with the concept of meaning and its role in the learning process. Not unlike many current theorists, Titchener (1915) stated, "When one and the same experience has different meanings it is because the context varies...when we make a mistake in meaning, it is because we supply a context of our own" (p. 119). Although Titchener was not specifically addressing comprehension monitoring with this statement, its inclusion in a discussion intended to define what it means to comprehend, is appropriate. In many circumstances, comprehension is defined not by some rule, but by the context (i.e., the individual). One especially striking conclusion to be drawn from the following discussion, is that even when definitions of comprehension are limited to an academic setting, these definitions may vary enormously.

One explanation for this diversity has been presented by Markman (1981) in an especially insightful discussion of comprehension monitoring. She has suggested that when we speak about understanding something, we assume that first, the object of our understanding is composed of structured or organized information; and second, this understanding often involves inferential processing [*italics added*].

She points out, for example, that to say that one "understands" a phone number is probably not a legitimate use of the expression "to understand." Phone numbers are things that can be learned and recalled, but not really understood. Most evident, she suggests is that:

Understanding is not all or none; it admits of degrees...it may be difficult to specify any endpoint of perfect understanding [since]...there may not be clear criteria for perfect understanding (p. 65).

Despite its active, inferential, schema building nature, some activities intrinsic to the act of comprehension may not always be immediately conscious. Flavell (1981), for example, has suggested that conscious metacognitive experiences are more likely to occur in novel situations or situations that explicitly require it (e.g., complex problem solving). This awareness is also particularly obvious in at least some cases where comprehension fails. That is, for mature readers, much reading and comprehending appears to occur on "automatic pilot." Experienced readers may "operate with lazy processors" in the sense that their reading activities are so well learned and practiced that written information is almost automatically translated to meaning (Brown, 1980). When recognition of a comprehension failure occurs, it does so in the form of some triggering event. This event may be the realization that one has read pages on end with no awareness of what they were about, the inability to define a particular word or concept, or simply a feeling of uncertainty or confusion. These metacognitive experiences concerning comprehension, in turn affect processing and "reprocessing" strategies. This corrective action is also included within the framework of metacomprehension research.

Educators and researchers concerned with applications to learning and instruction, have frequently defined comprehension within the confines of taxonomies and classification systems. Many descriptions consist of a limited number of categories which can be used to classify knowledge according to the cognitive operations purportedly required to learn the information. The most well known of these, the Bloom Taxonomy (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956), divides the cognitive domain into six categories ordered from least to most complex. Comprehension is defined by Bloom et al. (1956) as "those objectives, behaviors or responses which represent an understanding of the literal message contained in a communication" (p. 89). The levels are hypothesized to represent a cumulative hierarchy of knowledge beginning with a classification referred to as "Knowledge" ("those behaviors and test situations which emphasize the remembering, either by recognition or recall, of ideas, material, or phenomena" p. 62) and ending with a classification defined as "Evaluation." In between, in ascending order of cognitive complexity are levels identified as comprehension, application, analysis and synthesis. One potential weakness in the cognitive-process approach to knowledge classification is that these processes are inferential constructs and cannot be observed directly. Bloom et al. (1956) have recognized this drawback by admitting that it is not always possible to know if a student's response to a test question originated through a higher-level reasoning process or was acquired by reading it in a textbook (i.e., knowledge recall).



Other classification systems have taken a relatively dichotomous approach in describing levels of knowledge or comprehension. Pearson and Johnson (1978) differentiate textually explicit knowledge from textually implicit knowledge. The former refers to information appearing directly or literally in the text. The latter, involving the use of inference, refers to information that is present in a less obvious form. Scriptally implicit knowledge, a third classification, points to the importance of the reader's script (schema) in obtaining meaning. Using a similar approach, Svensson (1977) has described holistic and atomistic types of knowledge. The atomist is described as primarily oriented to memorization, parts, and details of the text; in contrast, the holistic learner is seen as one who "steps back" in an attempt to focus on the overall meaning and relationships inherent in the material.

Additional relevant research and theory concerned with conceptualizations of knowledge, understanding, and comprehension have been postulated by several cognitive developmental theorists, most notably Piaget (1926), Kohlberg (1981) and Perry (1970). Although each of these theories focuses on a particular realm of intellectual or ethical development, common to all of the theories is the belief that cognitive development proceeds through a series of stages each of which is qualitatively distinct and increasingly more complex. Movement through the sequence occurs as significant changes and experiences within the person and the environment interact and result in alternating periods of equilibrium and disequilibrium.

Whereas Piaget essentially centered most of his observations on the childhood years, William Perry and his associates (1970) presented a cognitive stage framework focusing upon intellectual development in the college years. Working with interview data collected from Harvard-Radcliffe undergraduates, Perry devised a developmental scheme in which college students are described as moving through a series of cognitive positions characterized as representing first dualistic, then relativistic, points of view. The dualist conceives of knowledge as quantitative and fact oriented; upon reaching relativism, knowledge is viewed as qualitative and dependent on context. Students in different stages are believed to view "knowledge" and "learning," the role of the teacher, and one's role as a learner in different ways and therefore, to respond differently to academic tasks present in their learning environment (Knefelkamp, Widick & Stroad, 1976).

How might individuals actually acquire particular levels of knowledge? Several encoding theories believed to account for the various levels of knowledge acquisition have been postulated during the past several decades. Craik and Lockhart's (1972) levels of processing theory conceives of memory as a byproduct of cognitive analyses such that the nature of a memory trace is determined by the level or depth of its processing. In text comprehension research, this has been translated to mean that subjects who employ "deeper" or more cognitively sophisticated monitoring strategies will retain more text information at all knowledge levels than those employing more "superficial," lower-level strategies (Kunen et al., 1984).

Although a number of studies have been shown to support the levels of processing framework, Morris, Bransford and Franks (1977) have suggested an alternative theory, the transfer appropriate processing model, to account for the interaction between the level of processing and the reader's learning goals. They argue that it is too simplistic to assume that some levels of processing are inherently more superficial than others. Instead, they counter that assumptions about the usefulness or inferiority of particular types of acquisition activities can only be made in relation to the existing testing situation. Therefore, although different levels of processing result in different types of information acquisition, all information may be equally strong and durable assuming that it is tested appropriately. For example, they suggest that if the criterion task is knowing the number of words in a list that contains "es," then "es-checking" is more likely to result in better performance than focusing on the semantic meaning of the words.

Although theorists frequently emphasize the active, inferential nature of comprehension, it can be noted that all of the previously described knowledge taxonomies reserve a place in their systems for a category of knowledge that would not ordinarily require much inferencing. Whether this level is referred to as dualistic, factual, textually explicit, or atomistic, its primary characteristic is that it is a "low level" or "preliminary level" of understanding (Bloom et al., 1956; Perry, 1970). Indeed, some have argued that such levels don't even meet the criteria for understanding. As Ornell (1979) has suggested, "true" understanding is distinct from both "knowledge" and "comprehension" [his terms]. True understanding goes beyond a taxonomy

of learning objectives or beyond simply knowing a collection of facts. Someone can demonstrate that they know something by having memorized it, but should that knowledge be considered understanding? Without becoming tangled in a word game, perhaps the best way to elucidate upon these thoughts is with reference to Markman's (1981) distinction between knowing and understanding phone numbers. (It makes sense to know a phone number but not to understand one). It should be pointed out, however, that in academic settings, A's can be acquired through knowing, understanding or comprehending--it's not the term that's important, it's the activity.

Teachers supply a definition of comprehension every time they construct a test. A useful and comprehensive definition of comprehension by an educator is that of Holt (1964). Making implicit reference to technically exacting terms like inferencing and higher level encoding, Holt's (1964) definition of comprehension includes the ability to "state it in your own words," "give examples of it," "recognize it in various guises and circumstances," and "foresee some of its consequences." Although educators frequently extol the virtues of these activities--at least in theory--in practice, extensive research on teacher's test questions and classroom tasks indicates that a large majority of these activities at the high school level encourage thinking at the lowest taxonomic levels (Gall, 1970; Trachtenberg, 1974). As Gall (1970) has stated, "Educators generally agree that teachers should emphasize the development of...critical thinking...Yet research spanning more than a half-century indicates that teachers' questions have emphasized facts" (p. 712).

Theorists have speculated about the effect of such experiences on student's own approaches to learning. Perry (1970) predicts that students will arrive at college at the dualistic, concrete end of the intellectual continuum, in part resulting from their particular stage of cognitive development. As a result of the diversity of ideas present in the college setting--and expected cognitive growth--these students come to challenge their original "black-white" thinking in favor of a more relativistic position. This evolution is seemingly demonstrated by a student's firsthand account of his learning style "then" and "now."

Then it was just the weight of the thing. Now it's, it's not so much how many pages there are on the reading list, it's more what the books are worth...I've finally decided that you don't read a book just to say you've read it...(Perry, 1981, pp 102-103)

In describing the apparent difficulty many college students have in adopting a holistic approach, Svensson (1977) has suggested that experience with factually oriented tests in secondary school may be partly responsible:

It may be difficult for [students] to realize that there are, or at least should be, different demands made in higher education. The development of personal knowledge and critical thinking are commonly thought to be an essential outcome of higher education...It seems essential to help students who adopt atomistic approaches to recognize the importance of interacting actively and critically with the information and ideas presented in the course of higher education (p. 243).

## CHAPTER 2

### REVIEW OF THE LITERATURE

#### Models of Learning

Two models of learning and metacomprehension, Flavell and Wellman (1977), and Jenkins (1979), serve as a useful starting point for organizing and interpreting the following literature review. Both frameworks point to the importance of considering multiple variables and their interactions in analyzing any learning activity. In the Flavell and Wellman (1977) model presented previously, four classes of phenomena--metacognitive knowledge, metacognitive experiences, cognitive goals and cognitive actions--were introduced as potential sources of influence in a cognitive enterprise. Three subcategories of metacognitive knowledge have been further identified. "Person" variables include knowledge of oneself and others as cognitive processors. "Task" variables concern available information about task demands and goals while "strategy" variables consider knowledge one might have about effective strategy use (Flavell & Wellman, 1977).

It is most practical to consider these variables in terms of their interrelations. For example, an individual might believe that unlike many of her friends her academic ability is best evaluated by essay tests (person variable). Furthermore, when she studies for essay tests she rewrites all of her class and text notes (person-task-strategy

interaction); in contrast, when she studies for multiple choice tests she has found that it is more effective to use flashcards.

Jenkins (1979) has called his model the "Problem Pyramid" or the "Theorist's Tetrahedron." Like the Flavell and Wellman (1977) model, it serves as a reminder of possible interrelationships among variables rather than a prescription for research design. The tetrahedron is presented in the form of a four-sided figure with vertices designated "subjects," "criterial tasks," "materials," and "orienting tasks." Edges of the tetrahedron represent two-way interactions, planes represent three-way interactions and the entire figure represents the four-way interaction. As Jenkins (1979) notes, in practice most researchers focus on one or two of their "favorite" vertices in their investigations. For example, a levels-of-processing theorist may be most interested in observing the effects of different orienting tasks while a developmental researcher might present the same task to subjects of varying ages. In each case, important knowledge about a cognitive activity is acquired. Yet, as Jenkins (1979) points out, "our statements about learning and memory must be couched in the general form of a four-variable relation: 'If the Subject...and if the Orienting Task...and if the Material...and if the Criterial Task..., then....'" (Jenkins, 1979, p. 441).

If, for example, a researcher collects data examining children's ability to recognize ambiguous messages, this may tell us little about adults' abilities to perform a similar task. The kind of task presented and the type of material used will be reflected in the learning activities of the subject. Some researchers have argued that many

studies have focused on either isolated task variables such as question types, lookback conditions and question formats or isolated learner variables such as prior knowledge or cognitive style (Davey & Lasasso, 1984; Kendall, Mason & Hunter, 1980; Samuels & Kamil, 1984). Changes in any of these major variables can alter the results of a study.

For example, in a study of word recognition, one can use an oral response, semantic categorization, lexical decision, or a matching task...[O]ther changes in addition to task and individual difference factors can alter the results of a study. The researcher may use high- or low-frequency words or the words may be presented in context or in isolation...A student who knows she is in an experiment as a requirement to complete a college psychology course may not react to the task in the same way as a businessman reading the morning paper on a commuter train (Samuels & Kamil, 1984, p. 190).

Although it has been argued that research failing to take into consideration these interactions creates serious limitations for many current findings, it must be recognized that the alternative can create hopelessly complicated research designs. A look at any contemporary discussion of the reading process illustrates clearly that these variables are certainly not being ignored (e.g., Pearson, 1984; Spiro, Bruce & Brewer, 1980).

The organizational scheme used in the following literature review is modeled after the framework provided by Jenkins (1979). Since many studies have examined the interactions of variables both within and among the four major components, any classification of a particular study is necessarily conditional. In addition, although a current trend in the metacognitive literature has been the identification and description of compensatory strategies, such a discussion is outside of the general interest of the present research and will not be included.



### Subject Characteristics

A large portion of the metacognitive literature is limited to young children. Brown (1980) has suggested that "one outcome of this limited focus is that a review of the literature could give the impression that metacognitive development is rapid and functionally complete by third grade" (p. 474). Indeed, as she points out in an updated review of the metacognition literature,

It is unfortunate that there is not more research activity in the area of adult metacognition. Anyone who has ever taught a group of college students must know that their metacognitive skills in a variety of domains could stand considerable enhancing! (Baker & Brown, 1984b, p. 380).

Two "subject" variables frequently identified as important in the metacomprehension literature are age and reading ability. Most studies generally conclude that younger and poorer readers are less effective in monitoring their comprehension than older and better readers. In her now classic research, Markman (1977, 1979) demonstrated that young school children are generally poor at analyzing oral messages for clarity and completeness. Children listening to a set of instructions on how to play a game and children listening to short essays containing inconsistent information, had difficulty reporting the inconsistencies or inadequacies present in the message. The ability to identify these problems was shown to be dependent on age, the nature of the materials (either a set of instructions or a short essay) and whether children were alerted to the presence of the inconsistencies. Similarly, descriptions of oral communicative adequacy (e.g., referential communication studies) have also pointed to young children's difficulty

both in providing nonegocentric messages and in analyzing the messages provided by others (Cosgrove & Patterson, 1977; Glucksberg, Krauss & Weisberg, 1966; Piaget, 1926).

Younger readers have also been shown to experience difficulties with active strategic attempts to adjust their reading activities in response to task instructions, to identify the main ideas in a written passage, and to estimate their readiness for a test (Brown, 1980; Brown & Smiley, 1977, 1978). Predictably when the abilities of children who are also poor readers are assessed, they are found to apply metacognitive strategies less frequently than skilled readers (Short & Ryan, 1984).

Investigations of monitoring strategies used by good and poor college student readers have found that good readers report using more strategies more often than poor readers and are more flexible in their ability to read for various purposes. Differences between good and poor readers may be more obvious when reading high-level technical passages than when reading less difficult passages (Hare & Pulliam, 1980; Hare, 1981; Smith, 1967). It is not clear whether these differences in strategy use can be accounted for by a lack of knowledge about strategies on the part of poor readers, or difficulty in engaging these strategies when appropriate (Hare, 1981).

#### Orienting Tasks

Research attempting to discern distinctive learning styles among students often relies on descriptions of strategies (what the learner does to attain comprehension) to make these classifications. Asking

readers to describe what they know about their own comprehension monitoring behavior may be accomplished through the use of interview studies, in conjunction with a strategy use questionnaire, or concurrently with the reading or studying task. Strategies can be spontaneous (reader-based) or task-structured and may refer to either behavioral or cognitive activities. The following discussion will focus primarily on what learners do when they study as opposed to what learners do when their reading activities are concerned with immediate comprehension only.

A learner's comprehension goals may be expected to directly influence subsequent use of information processing strategies. Although neither the Bloom Taxonomy or Perry's (1970) dualism-relativism scheme focus on actual text monitoring behaviors, certain predictions about strategy use can be made with regard to these systems of knowledge classification.

The relationship between general epistemological orientation and use of comprehension monitoring strategies has been investigated by Ryan (1984) and Svensson (1977) and has led to specific predictions concerning knowledge orientation and academic success. Ryan (1984) has reported that students classified as dualists are significantly more likely to use learning strategies characteristic of Bloom's (1956) level one knowledge category. Correspondingly, comprehension criteria reported by relativists more often involved the comprehension or application categories of the Taxonomy. This latter group of students subsequently earned higher course grades in an introductory psychology course than students identified as predominantly dualistic in

orientation. In addition, it was observed that those students reporting multiple comprehension criteria were more successful than students reporting fewer or less varied criteria suggesting that both the nature and number of comprehension criteria can be predictive of course grades.

Similarly, Svensson (1977) noted that students classified as atomistic in their learning activities were less successful in their studies than students utilizing holistic cognitive strategies. In this case, approaches emphasizing activities such as remembering sentences and details and memorizing parts of the text (strategies which would be described as dualistic) were again associated with poorer performance than were strategies focusing on applications of the message to a wider context.

Strategy use in experimental settings is frequently observed under conditions in which learners remain generally uninformed about the criterial task. They may be told about the general format of the test, that the task will assess their comprehension of the material, or that they will be "responsible" for all the material presented to them. It might appear from such limited instructions that the researcher is assuming that the learner will approach the reading task with the same learning goals implicitly held by the experimenter (Garner & Anderson, 1981-82). That the unspoken understanding is that learning occurs in one of two ways--either you learn it or you don't (and somehow you should know how to learn it). The common practice in some journal publications of neglecting to report the actual instructions given to subjects appears to confirm this underlying philosophy. Without specific information concerning the nature of the criterion task it is

up to the learner to determine the probable goals of the studying activity and select strategies appropriately.

When subjects are asked to describe "what they did" while reading a passage, their answers exhibit both individual differences and certain commonalities in their reported cognitive processes. Responses will also vary depending on the method of reporting. Subjects instructed to "write down everything you remember doing and thinking while you read the passage" may report different activities than those asked "Did you use strategy X?" These issues, as well as a more fundamental concern with the learner's ability to accurately describe what she knows about her thinking have been consistently debated in the literature (Afflerbach & Johnston, 1984; Garner, 1982; Phifer & Glover, 1982).

When subjects are not cued to particular strategic behaviors, their reported reading activities as described in the literature, often seem to fall into approximately seven to fifteen categories (Alvermann & Ratekin, 1982; Garner, 1982; Garner & Alexander, 1982; Olshavsky, 1976-77; Ryan, 1984). Typically reported activities include rereading, skimming, underlining, personal identification, reading for details, and reading for main ideas. When a strategy use questionnaire is utilized, even more activities may be identified (Cioffi, 1986).

Although a number of activities appear to be available while studying, among those most frequently researched are notetaking, underlining, and rereading. Underlining has generally been reported to be the most popular study-related activity although its efficacy is unclear (Anderson & Armbruster, 1984). Theoretically, activities that

provide opportunities for the learner to interact actively with the material--notetaking, summarizing, student questioning, outlining--would be expected to be more effective (Anderson & Armbruster, 1984; Andre & Anderson, 1978-79; Orlando & Hayward, 1978). Findings from much of the studying literature, however, are difficult to interpret given the difficulty of matching behavioral descriptions of an activity with actual level of encoding. Much research indicates, as well, that students have a number of monitoring strategies available to them and that strategy use is highly individualized (Smith, 1982; Wagoner, 1984). Not surprisingly, the most conclusive finding appears to be that "the more specific the knowledge about the criterion event, the greater the effectiveness of studying" (Anderson & Armbruster, 1984, p. 660).

#### Criteria Tasks

The previous research has reported some of the findings relevant to studying strategies when knowledge of the criterion task is limited or not analyzed as an independent variable. Information about the use of monitoring strategies alone, however, would appear to be limited in the extent to which it contributes to an understanding of the comprehension mismatch experience. This experience considers both the nature of the learner's study activities and the nature of the criterion task with which learning is evaluated.

Almost any measure of comprehension raises concerns regarding validity, reliability and generalizability to other materials or tasks (Gibson & Levin, 1975). Baker and Brown (1984a) have recently reviewed many of the methodological approaches used in comprehension monitoring research, providing a useful analysis of the particular benefits and

shortcomings of each method. Perhaps the most straightforward approach is simply to ask subjects if or how well they feel they understand the material. Similarly, subjects may be instructed to study material until they feel they understand it, and are then administered a comprehension test. Although either of these methods would seem to have ecological validity--students probably rely on their feelings to some extent to gauge their studying progress--one difficulty in using a comprehension test is that it may confound memory with comprehension. In particular, when considering studies occurring in a laboratory setting, it is not always evident whether poor performance on a comprehension test is the result of ineffectual monitoring, inaccurate memory, or lack of motivation to approach the task in the same way a subject would prepare for a classroom test.

Several approaches for measuring comprehension rely on the subject's ability to process text which has been altered in some way. In the cloze technique, subjects are asked to use contextual information in order to supply words or phrases deleted from the passage. In text disruption paradigms, confusing or incorrect information is introduced into a text; subjects are presumed to be good at comprehension monitoring if they can identify the errors or problems (Baker, 1979; Baker & Anderson, 1982; Markman, 1977, 1979). Failure to report text confusions, however, has not been shown to be a very sensitive indication of poor comprehension. Subjects may use inferencing to resolve their comprehension difficulties. Subjects may be unwilling to admit to problems in the passage or in comprehending, particularly when it is seen as a challenge to an "expert's" point of view. The task

itself appears to lack ecological validity since subjects have no reason to expect deliberate text inaccuracy based on previous text-processing experiences (Baker & Brown, 1984a, 1984b).

Other approaches for assessing comprehension include open-ended recall questions, instructing subjects to write down everything they remember from the text, essay questions, recognition tests of various types, maze tasks (similar to the cloze technique), problem solving tasks, and on-line processing measures. Not unexpectedly, research has demonstrated that the choice of a particular testing procedure does not always allow generalizations to other operational definitions of comprehension (Kendall et al., 1980).

A small number of studies have observed subjects' abilities to modify their reading and studying behaviors in response to instructions or information specifying the nature of the criterion task. Although it makes intuitive sense to presume that students' studying behavior and learning outcomes will be more successful when they are given pertinent information concerning the task (Anderson & Armbruster, 1984), in actual classroom settings there is disagreement concerning both the extent and nature of such information.

Some instructors follow the unspoken "learn by taking the test" philosophy and presume that providing information about study goals in advance "gives the whole test away." These instructors may believe that students will learn what they need to do by taking the test; it is not entirely clear, however, whether students can or will effectively learn from the negative environment which accompanies their test-taking



mistakes (Bransford, 1979). Some educators have argued that it is unfair to require students to guess their instructors' definitions of comprehension.

The purpose of education is to effect change...Once we are clear about our aims and objectives, it is important to let our students in on the secret...we have no right to change other people without first letting them know what we are attempting (Stanton, 1978, p. 75).

Still others worry that when teachers assume too much of the responsibility in guiding students toward some goal, active reading and self-questioning may be discouraged (Baker, 1979). Indeed, students express disagreement among themselves regarding the amount of information that should be given prior to an examination (Kraus, 1984). At issue for those investigating the comprehension mismatch experience, is the relative importance of operationalizing definitions of comprehension for students. If students are aware of the particular comprehension goals required for a test, their studying is predicted to be more effective. Yet, it is also true that knowledge of the criterion task will not in itself affect performance unless it also changes studying behavior. Empirical research assessing the effect of prior information on studying and task performance has reported inconsistent findings.

Some research has reported that students' comprehension criteria may be modified by specific instructions presented prior to or during the studying situation. Garner and Alexander (1981) demonstrated that undergraduates provided with specific information pertaining to performance goals will set higher learning criteria for themselves and

spend more time studying than students informed only that they should read a passage in preparation for a subsequent comprehension test. These findings may suggest that at least in experimental settings, subjects may operate with a less demanding definition of comprehension unless task demands specifically encourage them to do otherwise (Kunen et al., 1984).

Other research has modified the nature of the instructions presented to subjects in order to observe the extent to which subjects will alter their methods of study in association with perceived task criteria. For example, when instructed to read a passage for details or for general impressions, good readers have been shown to adjust their strategies to the two purposes more flexibly than poor readers and perform more successfully on subsequent tests of comprehension requiring recall of details and general ideas (Hare, 1981; Hare & Pulliam, 1980; Smith, 1967). Achievement on comprehension tests has also been shown to vary as a function of instructions, question types and their interaction. Fredericksen (1975) presented groups of undergraduates with an essay describing the political, economic, and social conditions of an imaginary island. Subjects were told that they would be responsible for the content of the passage, the content of the passage and solutions to problems raised in the passage, or solving problems described in the passage only. As predicted, subjects' ability to recall the content of the essay and solve problems raised by the essay varied significantly with regard to the instructions. The "problem solving instructions only" group demonstrated significantly more inferential processing than either of the other groups but was poorest

at recalling details from the passage. The other two groups were comparable in their recall ability but differed significantly in their problem solving activity.

Similarly, research conducted by Postman and Senders (1946) demonstrated that college student subjects provided with one of six different types of reading instructions, differed significantly with regard to the kinds of information recalled from the passage. For example, subjects informed that they would be tested for details were more successful in answering questions relating to details than subjects instructed to read for purposes of timing. All students however, irrespective of condition, performed well on the "general comprehension questions," suggesting that most college student readers will recognize that certain information is likely to be essential to a text-processing task.

Since the 1930's when multiple choice tests began to receive classroom use, educators have expressed concern that students studying for multiple choice exams would emphasize memorization and recall over more "desirable" (i.e., "higher level") forms of knowledge (Anderson & Armbruster, 1984). Although the common sense belief--that essay tests in particular, encourage greater comprehension and organization of the material to be learned--remains strong, findings from empirical research have been inconsistent. That is, studies investigating the effects of anticipating one test type over another often report no differences in studying behavior and test performance for students expecting one kind of examination versus another.

Hakstian (1971) investigated the effect of test type expectation on subjects' study methods and test performance in both an ordinary classroom setting and a more highly controlled laboratory situation. Under both conditions, subjects were led to believe that they would receive either a multiple choice, objective, or combined multiple choice and objective exam. The actual comprehension measure was composed of both types of questions and assessed for knowledge at four levels of the Bloom Taxonomy. In both studies it was observed that amount of study time, studying behavior and degree of stress on facts, ideas, and applications, did not differ with respect to test type anticipation. Interestingly, scores on the objective items were negatively correlated with both degree of stress on facts and degree of stress on applications; as predicted, scores on the essay test were significantly correlated with degree of stress on applications. In his discussion of these results Hakstian (1971) suggests that for both studies "the outstanding feature...is that the type of examination anticipated appears to make very little, if any, difference to time spent, breakdown of material, or techniques employed in preparation" (p. 323).

Somewhat different findings were observed in Sax and Collet's (1968) manipulation of test type expectation. Students in an actual classroom setting were informed that they would be tested using either a multiple choice or a short answer format. Of the students receiving the multiple choice test, those who expected this test performed significantly better than those expecting the short answer test. There were no differences in performance on the short answer test between those who expected it and those who did not. These findings suggest

that knowledge of the test type per se may not be sufficient information to affect test performance; knowledge of the kinds of information tested is also important. The students had been informed initially that all items on both tests would stress the interpretation and application of principles. The investigators suggest that, contrary to common opinion, multiple choice items requiring applications and interpretations of information may be more difficult than an essay test requiring the same information since the former may also require the ability to make fine discriminations between similar answers. Thus, in this instance, subjects expecting the multiple choice test were perceived to have learned the material more thoroughly than those expecting the short answer test.

The effects of manipulating both test type expectation and studying instructions have been investigated in several studies with varying findings. Kulhavy, Dyer and Silver (1975) reported that high school subjects assigned to one of nine conditions in which both test mode and study instructions were varied, performed similarly on the criterion task irrespective of type of test expected. Rickards and Friedman (1978) also observed that total test performance did not appear to be affected by test mode expectancy in their investigation; however, examination of the kinds of information studied by subjects varied by condition. Subjects expecting an essay test took more notes on high structural importance (SI) information than subjects expecting the multiple choice test. Differences in notetaking appeared to carry over into test performance; students expecting the multiple choice test recalled more low SI items from the total test than students expecting

the essay test, and vice versa for high SI items. The researchers noted that the nature of the criterion task appeared to mask performance differences which became evident once the individual criterion items were examined.

Kunen et al. (1981) have suggested that research designs administering comprehension questions on a poststudy basis, often bias the subjects' studying and encoding activities in favor of the lowest taxonomic levels (i.e., Bloom's level one category). When students are not provided with appropriate comprehension goals, they argue, they are likely to encode information at the level that is generally taught in school--according to much published research, these are the lowest cognitive levels (Gall, 1970; Trachtenberg, 1974).

In their investigation, undergraduate subjects were assigned to one of four taxonomic category treatment groups: knowledge, application, synthesis, and evaluation. The material to be learned consisted of sets of principles from various academic fields, altered slightly for each "orienting" condition. Subjects in each group were presented with a set of questions or tasks, all covering similar information, and differing only in the level of knowledge required. Following these tasks, all subjects were given an unexpected recall test containing many of the items appearing in their original orienting material. This memory task was considered to reflect knowledge of the material at the lowest taxonomic level.

The findings indicated that prior taxonomic experience affected recall performance in the following order from most words recalled to least: synthesis, application, evaluation, and knowledge. Although evaluation is theoretically considered to be the highest of these levels, it did not produce the highest recall. The other levels, however, did exhibit the predicted hierarchical relationship. Kunen et al. (1984) have used these findings to claim moderate support for the cumulative hierarchical assumption made in the Taxonomy.

The previous findings point out that, although knowledge of the criterion task is theoretically expected to affect performance, it will not unless individuals change their study strategies accordingly. Some investigators have suggested that most studies reporting performance effects for test mode expectancy have employed "low meaningful materials" such as serial lists and paired associates as criterion tasks rather than "meaningful" prose passages (Kulhavy et al., 1975). Anderson and Armbruster (1984) have discussed several circumstances in which students might choose not to change their regular study behaviors in response to specific instructions. In some settings (e.g., laboratory experiments), students may believe that the text to be mastered or the test to be administered is so short that the material can be learned by routine reading. Conversely, in some situations subjects may feel that the material to be learned is so extensive that no matter what the strategy employed, mastery of the material is unlikely. Students studying material in a laboratory setting may operate with less stringent comprehension criteria than students studying in a classroom setting. After reviewing a number of studies

examining student study behavior, Anderson (1980) summed up the existing literature by reporting that "students have a well-established study plan that is not easily modified by varying the task demands or types of study materials in the session" (p. 495).

For this reason, studies in which subjects are "instructed" to utilize particular study methods should be monitored closely for confounding variables. Many investigators do not appear to consider a subject's regular studying behavior before assigning him or her to a particular study condition. This practice can be disruptive for students who would not ordinarily employ such methods.

Finally, information about the testing or studying condition to which a subject is assigned does not necessarily reveal the actual nature of the encoding activities used by the subjects. Subjects who are assigned to a notetaking condition or report using a notetaking strategy, may differ in terms of the depth of their processing efforts. One "notetaking subject" could simply be copying material verbatim from the text; another "notetaking subject" could be reorganizing and applying the information to new contexts. Whether notetaking is shown to be effective will in large part be determined by the relationship between initial encoding activities and the requirements of the criterial task. It has been suggested that, for the most part, research has ignored the influence of the student's comprehension goals and the match or mismatch between his or her studying techniques and the comprehension goals reflected in the criterion measure (Anderson & Armbruster, 1984).



Of the research studies described in this section, only Hakstian (1971) and Rickards and Friedman (1978) collected reports of students' studying behavior and comprehension goals while preparing for a subsequent examination. Although a number of the other studies described here appeared to have created potential mismatch conditions by manipulating test type expectation and actual test received, without information about the learner's encoding processes it is difficult to define these testing situations as comprehension goal mismatches rather than some other kind of mismatch (e.g., test type mismatch).

A further controversy is reflected in the research findings reported by Kunen et al. (1981). Their findings provide partial support for the cumulative hierarchy assumption made in the Bloom Taxonomy by demonstrating that each successive component in the hierarchy represents the cumulative knowledge orientation contributions of the preceding levels. In their study, subjects "oriented" at level one of the Taxonomy--the knowledge level--were later tested with a level one recall test. Subjects in the knowledge condition were least successful on the test while subjects oriented at "higher" levels were more successful. These findings were interpreted within the levels of processing framework as support for the assumption that the "deeper" the level of processing, the greater the recall. This observation would appear to contradict the comprehension mismatch hypothesis by emphasizing the nature of the processing activities over the content of the criterion task.

In contrast, Fredericksen's (1975) findings indicate that subjects who performed least successfully on a recall test were those who were encouraged to be the most inferential in their processing (i.e., they were oriented to the passage at a "higher" knowledge level). Similarly, Rickards and Friedman (1978) observed that subjects who expected to receive an essay test, took more notes on high SI information and were subsequently more successful on the high SI items than the low SI items. The alternative was true for subjects expecting a multiple choice test. These latter findings appear to support the position advocated by Morris et al. (1977) by illustrating the importance of a match between initial processing and subsequent performance criteria. The confusing array of results can probably be reconciled to some extent by recognizing that true knowledge of a subject's processing activities is at best quite limited. In addition, actual comprehension tests may vary extensively and are not necessarily comparable (Gibson & Levin, 1975). Perhaps the most straightforward interpretation of the current state of the literature is that made by Anderson and Armbruster (1984) who suggested that studying is expected to be effective if students study the "right" information in the "right" way.

#### Characteristics of the Material

The final component in Jenkins' (1979) model of learning concerns the nature of the material which subjects use. Specifying the structure of a text allows a researcher to determine its generalizability to other passages, the amount and type of information which readers are likely to recall, and particular variables within the text which may affect the reader's understanding (Meyer & Rice, 1984). Some of the factors which

affect the amount recalled from a text include the organization of propositions and ideas, the number of concepts expressed, subjects' prior knowledge of the topic, vocabulary difficulty, sentence length, and reader interest (Goetz & Armbruster, 1980; Kintsch & Miller, 1984).

One traditional means of assessing text structure or difficulty has been through the use of readability formulas. Although they have not always been judged to be particularly sophisticated in their evaluation of the meaning and organization of a text, their popularity lies in the fact that they have been extensively validated and are straightforward in their application. Many current theorists point out that these measures typically consider only "surface" factors in the text such as sentence length and vocabulary density; a number of alternative models have been proposed more recently as a means of analyzing the complex interactions between text structure and reader characteristics (Meyer & Rice, 1984).

Although it is clear that the type of material presented to the subject determines the quality of comprehension, a detailed description of contemporary research on text structure is beyond the scope of the present literature review and research project. Descriptions of texts used in the studies described here, indicate that the following characteristics are most frequently taken into consideration: topic, word length, subject familiarity with the topic, readability, and reader interest.

### Overview of Research and Statement of Purpose

From this literature review it can be seen that research pertaining to learning and study monitoring has investigated a number of variables related to a final outcome called comprehension. Four major variables have been identified--subjects, criterial tasks, orienting tasks, and materials. They are described here primarily as a conceptual starting point for organizing an extensive body of literature which has as its goal the understanding and prediction of learning outcomes. Perhaps most challenging for the researcher attempting to contribute something original to this body of knowledge is the realization that there is no one "best story." As Jenkins (1979) has pointed out, "The overwhelming finding [from research examining these four variables] is that interactions abound and...straightforward generalizations about learning and memory are...rare" (p. 440).

The general findings from research considering subject differences and learning strategies have suggested that orientations toward and strategies for knowledge acquisition vary among college students (Perry, 1970; Ryan, 1984; Svensson, 1977). Distinctions have been made between students who use dualistic or factual approaches in their learning and students utilizing relativistic or applied learning strategies. These categorizations have been used to describe both learners and monitoring and study behaviors.

Studying has been described as reading in preparation for some criterion task. Specific knowledge about the task may vary extensively; under many ordinary learning circumstances students are likely to know

something about the task but do not have complete access to the task itself. Research investigating the relative effectiveness of varying study behaviors has generally concluded that as specific knowledge of the task increases, studying outcomes improve. There is, however, a large grey area in this prediction. That is, when some information about the task is known but not all information--as is usually the case--performance will vary in success depending on the specificity of the information. Furthermore, although knowledge of the criterion task may be seen as an important condition for optimal studying, it is not a sufficient condition. The relevant information must be processed appropriately and then reproduced or recalled in some fashion.

Several theoretical models have been introduced as a means of accounting for these encoding processes. One framework, the levels of processing theory, postulates that deeper levels of encoding will result in more durable memory traces. Kunen et al. (1981) successfully applied a levels of processing framework in their investigation of the Bloom Taxonomy. They have suggested that "lower" levels of knowledge are subsumed within this hierarchy such that subjects employing higher level processing efforts will be more successful on tasks requiring both high and low levels of knowledge.

A similar viewpoint is expressed in the work of both Ryan (1984) and Svensson (1977) who report that students described as holistic or relativistic in their knowledge orientations attain greater academic success than their dualistic or atomistic counterparts. Both researchers neglected, however, to fully describe the criterion task upon which academic success was based. It is not clear from these

findings whether the use of holistic or relativistic cognitive strategies is optimal for all knowledge levels of test items, or whether in some cases, strategies that rely on so-called less sophisticated encoding are also appropriate.

A modification of the levels of processing framework has contributed the concept of transfer appropriate processing as a means of understanding the interaction between encoding strategies and the goals and purposes of the learner. This model suggests that studying will be effective if the learner processes the "right" material in the "right" way. "Rightness" in this sense is related to the learner's knowledge of the criterion task and her ability to identify and focus attention on the relevant material. This notion of "rightness" also appears to be particularly applicable to match and mismatch experiences reported by students in classroom settings. The transfer appropriate processing model predicts, therefore, that certain studying or encoding strategies may actually impede learning in cases where "deeper, more meaningful processing" is not appropriate. Empirical evidence supporting this model has been provided by Rickards and Friedman (1978), Fredericksen (1975) and Morris et al. (1977).

Pilot data for the present investigation has suggested that one important explanation for a mismatch experience is that the subject feels she "knows something" about the material but determines after taking the examination that she didn't know it in the "right" way. This explanation has been interpreted by the present investigator as representative of the types of distinctions that have previously been made between knowledge orientations (e.g., Bloom et al., 1956; Perry,

1970; Svensson, 1977). That is, students may vary in their definitions of what it means to comprehend in an academic setting; such definitions may be predicted to match or mismatch with those represented in the criterion task. Little research has systematically examined mismatch experiences as they are described by college students. It further appears that researchers have largely ignored the influence of students' knowledge or beliefs about the criterion task and the relationship (match or mismatch) between the encoding strategies during studying and the requirements for performance on the criterion task. The following hypotheses will be tested:

1. It is predicted that a sample of college students will exhibit various levels of comprehension goals when presented with a reading-studying activity (Study One-Comprehension Goal Hypothesis).

2. It is predicted that differences in reported mismatch frequency will be related to general studying behaviors, comprehension goal ratings, comprehension of the passage, and reading ability (Study One-Mismatch Hypothesis).

1. Reading ability and consistency between comprehension goals and quiz type are important factors in quiz performance. Specifically, it is predicted that students receiving quizzes that "match" their comprehension goals will perform more successfully than students in comprehension mismatch conditions. The prediction favors the transfer appropriate model of encoding rather than a levels of processing framework (Study Two-Match-Mismatch Hypothesis).

2. It is further predicted on the basis of previously reported reading research, that several other variables will be influential in the prediction of quiz performance. Among these, reading proficiency has consistently been shown to be related to academic performance. Other variables expected to be important include reader interest, rated comprehension of the passage, dualism, and subjects' evaluations of their general mismatch experiences.

Subjects who describe themselves as experiencing frequent mismatches in the classroom are predicted to report more dualistic orientations than low mismatch subjects. This prediction is drawn from the general suggestion that college courses require more "applied" kinds of knowledge than "factual" (Perry, 1970), and from a specific observation by Ryan (1984) that relativists appear to have greater academic success in college than dualists (Study Two-Quiz Performance Hypothesis).

3. A validity check of the match and mismatch treatment is expected to reveal that post-quiz evaluations of the subjects' comprehension of the material and assessment of the quiz will be more negative in comprehension mismatch, as opposed to comprehension match, conditions (Study Two-Post-Quiz Evaluation Hypothesis).



## CHAPTER 3

### STUDY ONE

#### Method

##### Subjects

Subjects were 98 students from introductory psychology classes at a state university. 47 were male and 51 were female. The ages ranged from 17 to 19 with a mean of 18.2.

##### Measures

Stimulus passage. The stimulus passage consisted of a 3500-word passage taken from a psychology article (Goldberg, 1981). The passage appears in a book of readings in child development and was chosen to represent a typical college study (psychology) situation. The article has previously been used by the present researcher as part of the assigned reading for a college child development course. Briefly, the article compares and contrasts preterm and fullterm infant behavior and parent-infant interactions. A theoretical model based on previous research concerning learned helplessness and competence motivation is used to explain success or failure outcomes in these interactions. The passage was judged to be an interesting, though relatively unfamiliar topic for college students based on previous use in a classroom setting. Examination of the passage indicates that it presents a number of ideas, both factual and theoretical, which vary in importance and complexity.

The passage falls within the college level readability scale using the formula presented by Fry (1968).

Comprehension Goal Measures. The objective in designing a comprehension goal measure was to develop an instrument that would allow a subject to externalize information about his or her comprehension goals (reading purposes) in a context that would be sensitive to the qualitative differences that exist between different levels of knowledge (e.g., the Bloom Taxonomy). A 16-item comprehension instrument was developed for the stimulus passage. The measure was constructed as a traditional comprehension test might be; however, the purpose of the measure was to assess subjects' evaluations of each of the comprehension questions, not their ability to answer them. Since memory may be confounded with comprehension, subjects' ratings of the appropriateness of each question was perceived to be a more sensitive index of students' comprehension goals than their ability to answer the question correctly.

The questions in the instrument varied in textual importance and level of knowledge. Questions were classified using Bloom et al.'s (1956) criteria. Five questions were written to assess comprehension at the "knowledge" level (awareness, recognition or recall of certain facts or principles). Five questions assessed comprehension at the "comprehension" level (facts and principles can be restated or interpreted in another form). Five questions utilized a combined application-analysis comprehension criterion (principles can be applied in new situations and interrelationships are identified). The question formats followed those provided by Bloom et al. (1956) as closely as possible. One question in the test served as a distractor item. Two

versions of the comprehension instrument were constructed. One version required subjects to rate the appropriateness of each question and provide an open-ended reason for that rating. The slightly modified alternative version instructed subjects to both answer and rate each question.

Studying and Testing Experiences. A questionnaire was also constructed to assess subjects' testing, reading and studying experiences. In addition to demographic information, subjects provided information concerning their reading and studying behaviors, their perceived reading proficiency, positive and negative mismatch experiences, and interest in and knowledge about the preterm infant passage.

Mismatch was defined in the questionnaire as follows: "In general, how frequently have you gone into a test situation feeling like you understand the material, or that you studied as well as you could, only to be disappointed with your grade?" (A similar version assessed mismatches in the "positive" direction). Subjects responded by placing an X on a 7-point scale with the endpoints being almost never (1) and almost always (7). A list of nine possible reasons for experiencing a mismatch, previously collected from a pilot sample, was also included. Additional space was provided for subjects to write their own reasons if desired. Subjects were asked to select the three reasons that seemed to most frequently explain their own mismatch experiences.

## Procedure

The data were collected during group sessions of approximately 45 students each. All subjects were initially informed that they were participating in a study about reading comprehension in college students. Subjects were given a copy of the stimulus passage and instructed to read the material in the same way they would for a class assignment. They were informed that they would be asked to answer some questions about the reading following a 20-minute reading period. Subjects were permitted to make any marks on the passage they wished, but were otherwise not encouraged to use any specific reading strategy.

Immediately following the reading, the subjects received verbal instructions to rate their comprehension of the article on a scale ranging from (1) no understanding at all to (5) understood everything completely. Subjects then received one version of the "comprehension test" as an assessment of their comprehension goals while reading. All subjects within a single group session received the same measure.

The written instructions for the comprehension test asked the subjects to

read each question and think about how it corresponds to the reading; ask yourself, how fair or appropriate would this question be if it appeared on a test of this reading?

Subjects responded by placing an X on a 7-point scale with the anchors being very appropriate and very inappropriate. Following each rating, space was provided for subjects to indicate a reason for the rating. For the subjects required to both answer and rate the questions, the

additional instructions read "without looking back at the reading, answer the question as best you can." Subjects were given 25 minutes to complete this task.

The last task for all conditions was the Studying and Testing Experiences questionnaire. The entire procedure lasted approximately 70 minutes.

### Results

Preliminary Analyses. Demographic information about participants in the two conditions was analyzed to ascertain that the two groups were comparable. A gender (male, female) X response condition (rate and answer questions, rate questions only) analysis of variance showed that subjects in the two conditions did not differ significantly with regard to grade point average, interest in psychology, prior knowledge concerning preterm infants, experienced mismatch frequency, and most studying behaviors. Means and standard deviations are shown in Table 1. A significant main effect for condition was found in the comprehension ratings of the article. This difference favored subjects in the "rate only" condition who reported higher comprehension of the article than subjects in the "rate and answer" condition,  $F(1,59)=11.04$ ,  $p < .01$ . No other significant main effects for condition and no interactions were found.

Significant gender differences in interest in child psychology and interest in the article,  $F(1,59)=13.77$ ,  $p < .00$  and  $F(1,59)=7.29$ ,  $p < .01$  respectively, were found. In both cases, male subjects reported higher interest ratings than females. Significant gender differences

were also found with regard to three of the studying behaviors: highlighting,  $F(1,59)=4.47$ ,  $p < .05$ , outlining  $F(1,59)=4.61$ ,  $p < .05$  and summarizing  $F(1,59)=10.61$ ,  $p < .01$ . Females reported more frequent use of highlighting and outlining while males reported more use of summarization strategies. No other significant gender main effects were observed.

---

Insert Table 1 about here

---

Comprehension Goal Hypothesis. Appropriateness ratings for each of the questions comprising the three types of comprehension goals were computed by obtaining the mean rating for each of the levels--knowledge, comprehension, and application-analysis. The mean appropriateness rating for knowledge questions was 4.65 (.75), for comprehension questions 5.03 (.88), and for application-analysis questions 4.21 (.90). Eight different categories of reasons for the question appropriateness ratings were identified in subjects' written responses. These reasons are shown in Table 2 arranged by question level. A preliminary oneway ANOVA indicated that there were no significant differences for ratings by level for subjects who rated and answered the comprehension test questions and for subjects who only rated the questions,  $F(1,59)=1.54$ ,  $p > .21$ . Subsequent analyses combined subjects from the two conditions.

---

Insert Table 2 about here

---

A 3 X 2 repeated measures multivariate analysis with comprehension level (knowledge, comprehension, application-analysis) as the within subjects factor and mismatch (high frequency, low frequency) as the between subjects factor revealed a significant main effect for comprehension level  $F(2,59)=19.10$ ,  $p < .01$ . The main effect for

mismatch and the interaction between mismatch and comprehension level were not significant. Tukey post hoc comparison tests ( $p < .05$ ) on the differences between the means revealed that the difference between the comprehension level and the application-analysis level ratings was significant; the other comparisons failed to reach significance. Subjects rated questions at the comprehension level as most appropriate and the application-analysis level questions as least appropriate. Pearson product-moment correlations between the comprehension goal, mismatch, and comprehension ratings are reported in Table 3.

Insert Table 3 about here

Mismatch Hypothesis. In order to examine the issue of whether students with different mismatch frequencies differ in studying behaviors and comprehension goals, a oneway analysis of variance by mismatch level (high frequency, low frequency) was performed on the dependent measures. As shown in Table 4, subjects with mismatch frequency scores below 4.5 were classified as low mismatch subjects ( $n=47$ ) while those with scores of 4.5 and above were classified as high mismatch subjects ( $n=51$ ). One significant main effect for mismatch was observed. High mismatch subjects reported more prior knowledge than low mismatch subjects,  $F(1,97)=5.44$ ,  $p < .05$ .

Insert Table 4 about here

Pearson product-moment correlations, presented in Table 5, reveal that comprehension of the passage and self-rated reading ability were both negatively correlated with mismatch experiences. In both cases, the lower the comprehension rating for the passage and the lower the perceived reading ability, the more frequently subjects reported

experiencing mismatches. Mismatch also showed low but significant negative correlations with two of the studying behaviors: self-questioning and rewriting test material in one's own words.

Insert Table 5 about here

Nine different reasons purported to most frequently explain mismatch experiences had previously been identified in a pilot sample of undergraduates (Kraus, 1984) and are shown in Table 6. These explanations were presented to subjects in the present study who were instructed to select the three reasons that "most frequently explain your own [mismatch] experiences." Subjects were not told to rate the reasons in order of importance but only to select the three most common from the list.

Insert Table 6 about here

#### Brief Discussion

Study 1 was conducted for the purposes of obtaining an initial descriptive representation of college students' mismatch experiences. A comprehension test designed to reflect three levels of knowledge was presented to subjects in order to assess the frequency and comparability of these goals as well as their relationships with reported mismatch experiences.

Comprehension Goals. Results from Study 1 indicate that comprehension goals reflected in the quiz items differed significantly in their appropriateness ratings. Reference to Table 2 indicates that for all levels, the most frequent explanation for a question rating pertained to the extent to which the question was related to a "main



idea" or an idea believed to be important for overall understanding for the article. Examination of the explanations given for each level indicate that for "knowledge" questions, a high proportion of responses focused on issues related to overall importance or unimportance reflected by such comments as "too picky" or "not important for understanding." This is not surprising given the level one emphasis on recall and recognition of facts, in some cases perhaps--facts that are not believed to be central to an understanding of the passage.

Level two "comprehension" questions were most frequently rated as "main ideas," a reflection of the high appropriateness ratings given to this group of questions overall. Although level three questions also received a high number of "main idea" explanations, they were rated least appropriate in the overall ratings. This would appear to be explained by the extreme diversity of ratings for these questions. Although they were frequently rated as "main ideas," they were also frequently rated as inappropriate. Reasons given for inappropriateness ratings tended to suggest that some subjects believed that these questions covered information "not in the article" or raised issues that were too "opinion oriented." The diversity of these responses is not particularly surprising since level three questions are considered to represent a more complex integration of the text material. These questions were constructed to "go beyond" explicitly stated ideas; subjects were required to draw inferences and make applications. Although some students rated these questions as appropriate, others indicated that they were unfair primarily because they appeared to require answers that were arbitrary or entirely subjective. ("It's

based on my own opinion, how can you grade that?").

Although differences in comprehension goal ratings could not be attributed to general mismatch experiences, they do suggest that subjects view these comprehension goals as varying in appropriateness, at least within the confines of this particular experimental setting. Interestingly, level two questions were considered most appropriate, although level one questions have generally been considered easiest (Bloom et al., 1956). Since subjects' answers to the questions were not graded, it is not certain whether difficulty level is an important criterion for rating. Based on subjects' reasons for their ratings, however, difficulty level per se does not appear to be most important. Additional questions are raised regarding subjects' reasons for the question ratings. It is not clear, however, whether the ratings reflected the actual content of the question, the way that the information was presented, the knowledge level, or some combination of these. Since the three comprehension levels did not cover equivalent content areas it is not possible to draw firm conclusions concerning the rating differences.

Mismatch Experiences. Although reported mismatch frequency was not significantly related to question ratings and studying activities it seems reasonable to conclude that this experience is common to college students, at least the predominantly freshman group observed in Study 1. Fully 66% of subjects reported mismatch experiences in the medium range of the scale (between ratings of 3 and 5) with another 21% reporting mismatch experiences above that. Several general observations can be made about the reasons for mismatching presented in Table 6.

First, it appears that students are less likely to blame instructors or the test than might be imagined. Students appear to focus on their own abilities and activities as explanations for mismatching. Second, the two most frequently given responses appear to be interrelated: that the student feels she studied something but not what was "important to study" and that the student may have felt "I knew the information but not in the 'right' way." A substantial portion of the subjects reported experiencing test anxiety (40%) and many subjects (46%) considered studying proficiency as another important reason for their own mismatches. One interesting though not unexpected observation in regard to subjects' discussion of studying proficiency, was the finding that the most frequently engaged in study behavior was highlighting, an activity that has not generally been viewed as especially effective (Anderson & Armbruster, 1984). The next most frequently reported approaches were use of study guides or other text supplements and rereading. As suggested previously, study guides may not be particularly advantageous unless they encourage students to learn the same kind of information that will appear on their test. If they do not, it is possible that students may indeed be "learning the information but not in the right way."

The only significant findings related to the mismatch experience were with regard to two of the studying behaviors, self-rated reading ability and comprehension of the passage. Although the observed correlation with reading ability is not unexpected, it is evident that estimating one's own reading proficiency is a subjective process not necessarily independent from an estimation of one's mismatch frequency.

A more conclusive means of determining the importance of this relationship would be provided through the use of a standardized reading test.

Returning to one of the questions of the present investigation--"Why do mismatches occur?"--some initial hypotheses can be drawn from the present data. First, although mismatches have been attributed to various kinds of phenomena, most commonly they seem to be related to a situation in which the individual felt that she "knew" something but not in the "right" way. This interpretation lends support to the hypothesis that matching or mismatching in a test situation is a matter of learning or encoding material appropriately, where "appropriately" refers to the comprehension goals selected in the criterion measure. In addition to studying behavior, reading proficiency would be expected to play a role in the mismatch experience since previous observations report that better readers are more likely to vary their strategy use and read for "more purposes" than poor readers (Hare & Pulliam, 1980).

## CHAPTER 4

### STUDY TWO

#### Introduction

The main hypothesis explored in Study 2 was that students who receive quizzes expected to "match" reported comprehension goals will perform more successfully than students who receive quizzes that "mismatch" with their comprehension goals. Since a strong relationship between reading proficiency and academic performance has previously been established, a measure of reading ability will be included for use as a covariate in the analysis of quiz performance. Further exploration of the comprehension mismatch experience will be made possible through the use of an established reading test. It should be pointed out that the mismatch phenomenon will be studied from two perspectives in the present study. First, general mismatch experiences will again be explored as they relate to comprehension goals and reading ability. Second, an attempt will be made to manipulate subjects' match and mismatch experiences in the specific learning environment created by a reading and testing situation. This manipulation is expected to occur by random distribution of quizzes which are either factual or applied in knowledge orientation.

## Method

### Subjects

140 undergraduates from introductory psychology classes at a state university served as subjects. 66 were male (47%) and 74 were female (53%). The ages ranged from 17 to 32 with a mean of 19.1. 76% were freshmen and 19% were sophomores.

### Measures

Stimulus passage. The stimulus passage was a 2100-word condensed version of the passage described in Study 1 (Goldberg, 1981). The shortened version was used because subjects were allowed unlimited time to read the article.

Comprehension Goal Measures. A text-specific comprehension goal measure was constructed to assess subjects' knowledge orientations to the stimulus passage. The 16 statements represented both factual (8 statements) and applied (8 statements) types of information that could be learned from the passage and are shown in Appendix A. The statements were presented using the following format: "while reading the passage, how important for your understanding of what the article was about, was it for you to (comprehension goal statement)?"

The rationale for using a dual classification scheme (factual/applied), rather than the threefold system used in Study 1, was based on several considerations. First, this dual system of classification has previously been demonstrated by Ryan (1984) to be a successful means of classifying students according to their study strategies. Second, since Perry's (1970) dualism scale utilizes a dual

classification scheme previously shown to be related to the Bloom Taxonomy (used by Ryan, 1984), a twofold categorization was considered appropriate to maintain consistency among all measures.

Adherence Scale. The Adherence scale is a 7-item scale designed to assess students' knowledge orientations using a dualism/relativism framework. The scale used in the present study is a shortened version of Perry's (1970) scale, previously modified by Ryan (1984). Perry (1968) has shown each of these statements to load heavily on what has been described as an Adherence factor. Adherence is defined as a desire for "explicit and externally sanctioned structures of rightness" (quoted in Ryan, 1984, p. 250). The stronger the agreement with each statement, the more dualistic the individual. Ryan (1984) reports the test-retest reliability of the shortened version to be .84, comparable to that of .78 obtained by Perry (1968) for the full 46-item scale. The seven items are shown in Appendix B.

Comprehension Test. Two sets of multiple-choice questions were prepared to evaluate the subjects' understanding of the stimulus passage. One set of questions reflected the level one knowledge criterion described by Bloom et al. (1956) and a second set reflected a combined level two and three (comprehension-application) orientation. The combined level two and three classification has previously been recommended by Ryan (1984). Each question was also constructed to meet the textually explicit (i.e., level one) and textually implicit (level 2-3) criteria established by Pearson and Johnson (1978). To the extent that it was possible, questions for each version of the test covered equivalent content areas. Any questions which were not shown to meet

the criteria described above by two judges were discarded. The resulting comprehension tests were comprised of 12 questions each. As described in Study 1, each question was followed by a 7-point rating scale.

Nelson-Denny Reading Test (NDRT). Form D of the Nelson-Denny Reading Test (1973) was administered to subjects in order to assess reading proficiency and vocabulary knowledge. The NDRT is a timed examination measuring comprehension, vocabulary, and reading speed for students in grades 9 through 16. The examiner's manual reports a median correlation of .40 between scores on the NDRT and college achievement. Other studies have reported a mean correlation of .56 between the NDRT and semester grades in an introductory psychology course (Gerow & Murphy, 1980) and a correlation of .79 between verbal scores on the SAT and the NDRT (Erwin & Millikin, 1980). Because of its relationship to more general measures of achievement such as the SAT, researchers have recommended its use as a potential covariate in reading research (Zimmer, Glover, Ronning & Petersen, 1979).

Measures of comprehension and testing experiences. Subjects also completed the following measures during the course of the study: a) a "before quiz" and "after quiz" rating of understanding of the passage; b) a descriptive comparison pertaining to the pre-quiz and post-quiz ratings of comprehension; c) a rating of the perceived efficacy with which the quiz assessed the subjects' understanding of the passage; d) a rating of the subjects' perceived ability to judge test questions as fair or unfair; e) a "comfort" rating of the experimental setting; and f) ratings of mismatch frequency, interest in and familiarity with the



content of the stimulus passage and various demographic items. Several open-ended questions were included to provide subjects with the opportunity to discuss the reasons for these ratings.

### Procedure

Subjects attended two group sessions of approximately 25 students each. The sessions were separated by a 5-7 day interval. Since some subjects in Study 1 indicated that a quieter setting would have been more conducive to concentrated study, care was taken in the present study to select room conditions and groups sizes favorable to the task requirements.

During the first session subjects completed an informed consent letter which briefly described the study as an investigation of "reading and studying behaviors in college students." Before the stimulus passage was distributed, subjects were given the following verbal instructions:

During your reading of the passage you should read in the same way you would for a course assignment. You may make any marks you want on the passage. In addition, you should read at your own rate until you are satisfied with your comprehension. You may have as long as you wish to read and study the article. Later on during the session you will be given a multiple-choice quiz on the article. If you do not receive a basic minimum score on the quiz today, you will be asked to complete another quiz during the second session.

It was hypothesized that the instructions concerning the possibility of a second quiz, would increase subjects' motivation to complete the tasks conscientiously and treat the experiment seriously. Subjects signaled their completion of the reading by returning the passage to the experimenter. Subjects' reading times were recorded and they were given a packet of material containing the pre-quiz comprehension rating, the comprehension goal measures, and the

demographic questions. Their understanding of the passage prior to taking the quiz was rated in response to the following question: "Using the scale below, circle the number that best represents how well you feel you understand the article on preterm babies." The scale ranged from 1 (no understanding at all) to 7 (understood completely).

When all subjects had completed reading the passage and the items in the first packet of information, the comprehension quiz was distributed. Due to an unexpected problem in the quiz preparation, all subjects in the same group received the same quiz. The original plan called for random distribution of the two quizzes within each group session. Subjects were given 15 minutes to complete the quiz. The instructions for the quiz were as follows:

Following is a series of multiple-choice questions about the passage you read. For each question you should do the following: 1. Choose the one best answer for each question. Indicate your response by circling the appropriate letter. 2. Think about how the question corresponds to the reading and ask yourself how fair is it as a test question on this reading. 3. Following each question you will find a rating scale similar to the one below. For each question, circle the number that best represents how fair you feel the question is (as a test question on this reading).

The scale ranged from 1 (completely unfair) to 7 (completely fair).

All subjects completed the quiz prior to the time limit.

Following the quiz administration, subjects were asked to respond to a series of written questions regarding their experience and perceptions of the quiz. The first question concerned subjects' post-quiz comprehension. Subjects responded to the following question: "Using the scale below, circle the number that best represents how well you feel you understand the article on preterm babies after taking the

quiz." The scale ranged from 1 (no understanding at all) to 7 (understood completely).

Following this question, subjects were instructed to

place an X beside the one statement that best describes the comparison you could make between how well you felt you understood the article on preterms before taking the quiz and how well you felt you understood the article on preterms after taking the quiz.

The statements were presented in the following form: "After taking the quiz, I felt that my understanding of the article was [ ] than I had believed before taking the quiz." The adjectives inserted into the statements were: much better, somewhat better, a little better, about the same, a little worse, somewhat worse, and much worse. Subjects then described their reason(s) for their answer to this question.

This was followed by a question instructing the subject to "circle the number that best represents how well you feel the quiz measured your understanding of the article." The 7-point scale ranged from 1 (the quiz was a very poor measure) to 7 (the quiz was an excellent measure). Following this rating, subjects provided a reason for their response. The last question asked subjects to rate how well "you feel you are able to judge whether a quiz question is fair or unfair." The endpoints of the 7-point scale were 1 (I am very poor at making this kind of judgement) and 7 (I am very good at making this kind of judgement). The first session lasted approximately eighty minutes.

Subjects returned to a second session the following week. At the beginning of this session subjects completed a brief questionnaire assessing their general mismatch frequency and were asked to rate how comfortable they felt reading the article in the experimental setting.

The 7-point scale ranged from 1 (much more comfortable than my own study environment) to 7 (much less comfortable than my own study environment). Subjects were then asked to compare their high school study habits with their college study habits and rate their current level of satisfaction with their study skills. The scale ranged from 1 (not at all satisfied) to 7 (completely satisfied).

The final measure to be administered was the Nelson-Denny Reading Test (NDRT). The complete test was given to all subjects according to the instructions provided in the test manual (1973). Subjects were given 10 minutes to complete 100 vocabulary items and 20 minutes to read five selections and answer a series of multiple-choice questions at the end of each selection. While answering the questions, subjects were allowed to refer back to the passage. The first minute of this portion of the test was used to determine reading rate. Subjects began reading the first passage and at the end of one minute were told to stop on the line they were reading and record that line number on their answer form. The subjects then returned to their reading. Upon completion of the reading test subjects were informed of the purpose of the study during a 15-minute debriefing session.

### Results

Preliminary Analyses. Pearson product-moment correlations were computed to examine the intercorrelations between the various rating and performance measures and are reported in Table 7. Gender (male, female) X testing condition (factual, applied) analyses of variance were performed on the responses and test scores from the comprehension and testing experience questionnaire, the NDRT, the comprehension goal

measures, and the two quiz types to assess the comparability of the two quiz groups. Means and standard deviations are shown in Table 8.

---

Insert Tables 7 and 8 about here

---

The two-way ANOVA revealed a significant interaction between gender and quiz type on prequiz comprehension of the article,  $F(1,139)=8.14$ ,  $p < .01$  indicating that females in the factual group and males in the applied quiz group reported the highest levels of comprehension followed by females in the applied quiz group and males in the factual quiz group. The main effect for gender was also significant,  $F(1,139)=6.41$ ,  $p < .05$  with females rating their comprehension of the article higher than males. Since prequiz comprehension of the article was significantly correlated with both reading ability and interest in the article (see Table 7), these findings were reanalyzed employing both of these variables as covariates. This analysis again produced a significant interaction,  $F(1,139)=4.66$ ,  $p < .05$ ; no significant main effects were observed. A main effect for gender was observed with regard to the analysis of variance on interest in the article,  $F(1,139)=9.19$ ,  $p < .01$  when employing NDRT scores and pre-quiz comprehension as covariates. Females reported higher interest ratings than males. The main effect for quiz condition was not significant.

With regard to dualism scores, a main effect for gender was observed,  $F(1,139)=7.27$ ,  $p < .01$ , indicating that females were more relativistic than males. Since reading ability was moderately correlated with dualism (see Table 7) this analysis employed reading ability as a covariate. A significant interaction between gender and quiz condition was observed in the two-way ANOVA on prior knowledge of

preterm infants,  $F(1,139)=4.07$ ,  $p < .05$ . Females in the factual quiz condition reported the highest level of knowledge, followed by females in the applied quiz condition and males in the applied quiz condition. The main effects for gender and quiz condition were both significant,  $F(1,139)=8.43$  and  $F(1,139)=4.43$ ,  $p$ 's  $< .01$  and  $.05$  respectively. Females reported more prior knowledge than males while subjects in the factual quiz condition reported more prior knowledge than subjects in the applied quiz condition.

A significant quiz condition main effect for subjects' NDRT comprehension subtest scores was observed,  $F(1,139)=5.13$ ,  $p < .05$ . Subjects in the applied quiz group attained higher reading scores than subjects in the factual quiz group. This quiz condition difference was not observed, however, with regard to subjects' overall reading test scores,  $F(1,139)=3.07$ ,  $p > .05$ . The two-way ANOVA also showed a significant main effect for quiz condition on quiz scores when controlling for the effects of reader interest and reading ability,  $F(1,139)=59.60$ ,  $p < .00$ . Subjects in the factual quiz condition attained significantly higher scores on the quiz than subjects in the applied condition.

A chi-square analysis indicated that the two quiz conditions differed significantly with regard to gender,  $\chi^2(1, N=140)=5.88$ ,  $p < .05$ . Proportionately fewer males ( $n=24$ ) than females ( $n=41$ ) were in the applied quiz condition; this relationship was reversed for the factual quiz condition ( $n_s = 33$  and  $42$  respectively).

Factor Analysis and Reliability of Measures. For scoring purposes each subjects' factual, applied and dualism score was individually computed as the sum of the ratings comprising each of these measures. A factor analysis was performed to determine whether the statistical relationship of the items corresponded to the previously defined conceptual relationship. The factor analysis was performed on the twenty-three individual items composing these three scales with the commercially available SPSS program (Nie, Hull, Jenkins, Steinbrenner and Brent, 1975) using principal factoring with iteration and a varimax rotation. A four factor solution, shown in Table 9, yielded three primary factors which together accounted for 89% of the total variance.

The first factor (on which all of the dualism items loaded heavily) represents Perry's (1970) adherence scale. The second factor (on which three of the applied and three of factual comprehension goal statements load) represents the comprehension goals most closely associated with main ideas in the passage. The third factor (on which four factual and one applied comprehension statement load) consists of comprehension goals related to facts and details in the text. The remaining comprehension goal statements (one factual and four applied) did not load on any of these primary factors. Alpha coefficients for the three factors are .77, .72, and .62 respectively.

---

Insert Table 9 about here

---

Examination of the individual items comprising these factors indicates that the items loading most heavily on the second factor fall into two general categories: parent or infant behaviors and competence motivation. Items loading on the third factor contain comprehension

goals pertaining to definitions, facts and names. The results of the factor analysis suggest that the informational content of the comprehension goals was a more salient basis for rating than the nature of the cognitive processing. As a result of the factor analysis it was determined that the most consistent interpretation of these measures would be to consider them as main idea (factor two) and factual (factor three) comprehension goals rather than applied and factual comprehension goals. Subsequent analyses employed this distinction.

Match-Mismatch Hypothesis. The match-mismatch hypothesis predicted that subjects receiving quizzes which were expected to "match" their comprehension goals would perform more successfully than subjects in comprehension mismatch conditions. Strictly speaking, the restructuring of the comprehension goal measure into factual and main idea scales made a straightforward analysis of this hypothesis difficult. Furthermore, while examining subjects' main idea and factual comprehension goal orientations it was apparent that it would not be possible to classify subjects as either primarily factual or main idea oriented on the basis of their scores. The majority of subjects attained moderate to high scores on both scales simultaneously. On this basis, it was decided to utilize a median split procedure to classify subjects as high or low on each dimension. The resulting groups were further broken down into four subgroups comprising subjects representing one of the comprehension goal high and low combinations (high factual, high main ideas; low factual, high main ideas; low factual, low main ideas; high factual, low main ideas). Analyses of variance and Pearson product-moment correlations were computed to investigate the relationships between comprehension



goal status and quiz performance.

A main idea level (high, low) X factual level (high, low) analysis of variance on applied quiz scores found no significant main effects for either main idea,  $F(1,64)=.02$  or factual orientation,  $F(1,64)=.00$  ( $p$ 's  $> .87$ ). Using factual quiz scores, this same analysis also indicated that neither main idea orientation nor factual orientation was significantly related to quiz performance,  $F(1,74)=.40$  ( $p > .54$ ) and  $F(1,74)=1.31$ ,  $p > .26$  respectively. These findings indicate that subjects' reported comprehension goal strategies were unrelated to quiz performance irrespective of the type of quiz received.

Table 10 shows the correlations between factual (FA) and main idea (MI) comprehension goals, quiz scores (total, hard and easy halves), and NDRT scores for the four subgroups of subjects. Hard and easy halves for each quiz were derived by obtaining the percentage of subjects passing each item and combining the six questions with the highest, and six questions with the lowest percentage passing for each quiz type. This alternative analysis of quiz scores was used primarily as an exploratory technique in order to further examine whether the hypotheses required a qualification concerning the difficulty level of the two quizzes. Although the  $n$ s for each of the eight groups are small, it can be observed that factual and main idea comprehension goals are significantly correlated in several cases with the total or half-quiz scores. In particular, it appears that the relationship between comprehension goals and quiz performance varies considerably depending on the type of quiz received and subjects' comprehension goal classification.

---

Insert Table 10 about here

---

Quiz Performance Hypothesis. The quiz performance hypothesis predicted that in addition to the match or mismatch treatment, other factors including dualism, mismatch frequency, reading ability, and reader interest would be related to quiz performance. With regard to dualism, it has already been noted that degree of relativism or dualism was not significantly correlated with quiz performance once the effects of reading proficiency are partialled out. The correlation between dualism scores and reading ability reported in Table 7 indicates that as subjects become more relativistic in their knowledge orientations, reading scores increase. This hypothesis also predicted that subjects' reported mismatch frequency would be related to quiz performance with the expectation that subjects who reported experiencing frequent mismatch episodes would perform less successfully on the quiz than low mismatch subjects.

To examine differences among subjects reporting varying mismatch frequencies, subjects were classified into high, medium, and low mismatch groups. Subjects reporting mismatches less than half the time were considered to be low mismatch subjects ( $n=52$ ), approximately half the time--medium mismatch subjects ( $n=43$ ), and more than half the time--high mismatch subjects ( $n=45$ ). A quiz condition (2) by mismatch level (3) analysis of variance on the comprehension and performance measures produced one significant interaction with regard to time spent reading the article,  $F(2,137)=3.47$ ,  $p < .05$ . High mismatch subjects in the factual quiz condition reported the longest reading time, followed by moderate mismatch subjects in the factual condition and low mismatch

subjects in the applied condition. No significant main effects were observed. Means and standard deviations are reported in Table 11.

Insert Table 11 about here

Pearson product-moment correlations were computed to determine if the relationship between the various predictor variables and quiz performance differed for high, medium, and low mismatch subjects. Quiz scores for the two quiz conditions were converted to z-scores to facilitate comparisons across quiz type. As shown in Table 12, quiz performance for high mismatch subjects was significantly correlated with interest in the article,  $r(43) = .34$ , reading ability  $r(43) = .45$ , dualism level  $r(43) = -.29$ , and time spent reading the article  $r(43) = -.40$ . Quiz performance for medium and low mismatch subjects showed a somewhat different pattern. For medium mismatch subjects, quiz performance was significantly correlated with pre-quiz comprehension of the article,  $r(41) = .39$ , reading ability  $r(41) = .44$ , and factual comprehension goal orientation  $r(41) = .27$ . Quiz performance for low mismatch subjects was also significantly correlated with pre-quiz comprehension of the passage,  $r(50) = .36$ , and with interest in the article,  $r(50) = .47$ , reading ability  $r(50) = .28$  and dualism level  $r(50) = -.30$ .

Insert Table 12 about here

Several observations can be made about these data. The only variable consistently correlated with quiz performance across the three mismatch levels is reading ability. Pre-quiz comprehension of the article appeared to be predictive of quiz performance of medium and low mismatch subjects but not high mismatch subjects. Since all subjects were instructed to read the passage until they felt they understood it

well enough to take the quiz, this finding appears to suggest that high mismatch subjects' ratings of their comprehension were a poor estimate of their subsequent quiz performance.

Because the patterns of relationships between subjects' characteristics and quiz performance were not the same across mismatch groups, multiple regression analyses were performed to determine the best predictors of quiz performance. All cognitive measures and student characteristics served as predictor variables in the analysis; the criterion variable was the z-converted measure of quiz performance. As shown in Table 13, for students described as low in mismatch frequency, interest in the article and dualism level were the only significant predictors, yielding a multiple  $R$  of .47 and accounting for approximately 22% of the variance. Quiz performance was best predicted by reading ability, the pre-quiz comprehension rating and time spent reading the article in the medium mismatch group. The regression equation yielded a multiple  $R$  of .60, accounting for approximately 36% of the variance. Three variables also served as predictors of quiz performance in the high mismatch group. As with the medium mismatch group, reading ability was entered into the equation first, followed by subjects' ratings of their positive mismatch experiences and their interest in the article. The regression equation yielded a multiple  $R$  of .59, accounting for approximately 33% of the variance.

---

Insert Table 13 about here

---

Because the two quizzes differed in both difficulty level and knowledge orientation, independent analyses were conducted to determine the ability of the various comprehension goal and reading related measures to predict quiz performance. Stepwise multiple regression analyses were performed using reading ability, gender, comprehension ratings, interest ratings and comprehension goal items entered individually, as predictor variables. As shown in Table 14, pre-quiz comprehension of the article, reading ability, and six of the comprehension goal items were the best predictors for scores on the factual quiz. The regression equation yielded a multiple  $R$  of .71, accounting for approximately 52% of the variance. With regard to the applied quiz, reading ability and interest in the article were the only significant predictors, resulting in a multiple  $R$  of .53 and accounting for approximately 29% of the variance.

---

Insert Table 14 about here

---

Post-Quiz Evaluation Hypothesis. The post-quiz evaluation hypothesis predicted that subjects would rate the quiz more negatively in mismatch, as opposed to match, treatment conditions. Using subjects' quiz scores as a covariate, a factual level (high, low) X main idea level (high, low) X mismatch level (high, moderate, low) analysis of variance on subjects' post-quiz evaluations of the factual quiz revealed a significant two-way interaction between main idea level and mismatch frequency,  $F(2,138)=5.66$ ,  $p < .01$ . The most positive rating of the quiz was provided by subjects with a low main idea orientation reporting the highest mismatch frequency; The next most positive rating was provided by high main idea subjects with low mismatch frequency, followed by the

low and high main idea subjects reporting moderate mismatch frequency. The least positive rating was reported by subjects high in main idea orientation and high in mismatch frequency. No other significant interactions or main effects were observed.

Examination of subjects' mean ratings of the quiz reveal that for the low main idea group, as mismatch frequency increased, so did subjects' positive ratings of the quiz. The opposite relationship was observed with regard to high main idea subjects. The factual level X main idea level X mismatch level ANOVA for the applied quiz revealed no significant main effects or interactions. These findings suggest that subjects' evaluations of the testing experience were not related to their reported comprehension goals.

Gender (2) X quiz condition (2) analyses of variance using quiz score as a covariate were conducted to examine the effect of quiz condition on the post-quiz measures. The two way ANOVAs revealed statistically significant quiz condition differences for the post-quiz rating of comprehension  $F(1,138)=5.40$ ,  $p < .05$ , the comparison between pre- and post-quiz comprehension,  $F(1,138)=17.76$ ,  $p < .00$  and the overall evaluation of the quiz,  $F(1,138)=13.21$ ,  $p < .00$ . The main effects for gender and all interactions were nonsignificant. Subjects assigned to the applied quiz condition reported significantly lower postquiz comprehension, a greater discrepancy between pre- and post-quiz comprehension of the article, and a more negative evaluation of the quiz. The two way ANOVA also revealed a significant quiz condition difference in the mean ratings of the individual quiz questions,  $F(1,138)=11.59$ ,  $p < .01$  when controlling for quiz score with subjects in

the applied quiz condition reporting significantly lower ratings. The main effect for gender and the quiz condition by gender interaction were nonsignificant. Means and standard deviations for these variables are shown in Table 15.

---

Insert Table 15 about here

---

Although these measures were originally intended as validity checks of the comprehension goal mismatch manipulation, they also serve to suggest that overall, subjects assigned to the applied quiz condition experienced a mismatch in their expectations concerning the quiz. Examination of the means for the pre- and post-quiz comparisons, indicates that subjects in the factual quiz condition reported an overall rating of 3.69 while subjects in the applied condition reported a mean of 4.63. A rating of 4 on this scale indicates that the subject felt that the quiz was about what was expected; lower scores indicate that the subject underestimated her prequiz comprehension while higher scores indicate that the subject overestimated her prequiz comprehension.

Two observations should be made about this finding. First, subjects made their post-quiz evaluations without knowing their actual quiz score. Their mismatch ratings reflected their perception of their quiz performance. Second, it should be recalled that overall, subjects' reported pre-quiz comprehension was quite high and did not differ significantly with regard to quiz condition. Subjects were instructed to study the passage until they felt they understood it. Therefore, it is suggested that most subjects entered the quiz condition feeling like "they knew the information," an essential preliminary to the mismatch

experience.

Stepwise multiple regression analyses were also used to determine the optimal set of predictors for subjects' postquiz evaluations of the quiz. As indicated in Table 16, post-quiz comprehension of the article, dualism level and one of the comprehension goal items were the most significant predictors for subjects' evaluation of the factual quiz. In the case of the applied quiz, the best predictors were one of the comprehension goal items and subjects' ratings of their positive mismatch experiences. The resulting regression equations accounted for approximately 35% and 21% of the variance respectively. When pooling subject responses under both conditions, the most significant predictors of the postquiz evaluation were postquiz comprehension of the article, quiz type, dualism level, and the mean ratings for the individual quiz items. The stepwise multiple regression equation yielded a multiple R of .56, accounting for 32% of the variance.

---

Insert Table 16 about here

---

Responses to an openended question asking subjects to describe their reactions to the quiz were examined and response categories were derived which represented the criteria on which subjects based their evaluations. A summary table reporting the frequencies of subjects' mismatch ratings by quiz condition and the reasons for these ratings is presented in Table 17. While 23% of the students taking the factual quiz rated it as mismatched with their expectations, 62% of the subjects in the applied condition reported a mismatch experience. Examination of the reasons for these ratings indicates that student responses fell into three general categories: statements pertaining to quiz items, to the



quiz format, or to other aspects of the quiz or the learning-testing situation. The quiz item responses made reference to either the knowledge orientation of the items, the wording of the items or the fairness of the items.

Insert Table 17 about here

### Replication Study

A replication study was conducted following the analyses of the results in Study 2 due to the unequal gender distribution across the two quiz conditions. The primary purpose for this study was to observe whether the quiz score differences remained significant when controlling for the effects of gender and other subject characteristics.

### Method

#### Subjects

60 introductory psychology students with a mean age of 18.7 participated in the replication study. 30 subjects were male and 30 subjects were female. 83% of the subjects were freshmen.

#### Measures

All subjects received the stimulus passage described in Study 2, a brief questionnaire assessing interest in the article and prior knowledge of preterm infants, a 24-item quiz composed of all of the questions from the factual and applied quizzes arranged randomly, and a mismatch rating scale. Subjects did not complete the comprehension goal measures or the NDRT.

## Procedure

Subjects attended one group session, lasting approximately 90 minutes with 30 subjects scheduled per session. After completion of the informed consent letter, the stimulus passage was distributed. Subjects were given the same reading instructions described in Study 2 with the exception of the reference to a minimum quiz score. Following the reading session, subjects completed a questionnaire containing demographic information and the interest ratings. When all of the subjects had completed these materials the quiz was distributed. Subjects were read the same instructions as in Study 2 but were given 25 minutes to work on the quiz rather than 15 minutes. After the quiz, subjects completed the mismatch questions and attended a 15-minute debriefing session.

## Results

The primary concern of this analysis was to replicate the test score differences observed in Study 2. A two-tailed t-test comparison of applied and factual quiz items demonstrated that, consistent with Study 2, subjects performed significantly better on the factual items than the applied items,  $t(59)=7.21$ ,  $p < .00$ . A significant difference in the quiz item ratings was also found with regard to the factual and applied quiz questions. The factual items were rated as more fair than the applied items,  $t(59)=9.22$ ,  $p < .00$ .

A gender (2) by mismatch frequency (3) ANOVA for quiz scores, interest ratings and quiz item ratings revealed several significant main effects for gender but no main effects for mismatch experience. Female

subjects reported more interest in the article than males,  $F(1,59)=15.20$ ,  $p < .00$  and performed significantly better on the factual quiz items,  $F(1,59)=7.45$ ,  $p < .01$ . A corresponding main effect for gender with regard to the applied quiz items was not observed nor were there any significant gender by mismatch interactions.

Correlations were also computed to examine the relationships between the demographic and subject items and quiz performance. Quiz scores for the factual and applied quiz items were significantly correlated,  $r(58) = .47$ , as were the fairness ratings of the factual and applied items,  $r(58) = .55$ ,  $p$ 's  $< .001$ . The only significant correlations for the factual and applied quiz scores with regard to subject variables were with self-rated comprehension of the article,  $r(58) = .34$  and  $r(58) = .39$  respectively,  $p$ 's  $< .01$ . Quiz scores were not correlated with interest in the article as had been observed in Study 2. However, interest in the article was shown to be significantly correlated with prior knowledge,  $r(58) = .40$ ,  $p < .001$ . A significant negative correlation was again observed between self-rated reading ability and mismatch frequency,  $r(58) = -.28$ ,  $p < .05$ . No other significant correlations were observed.

In general, the findings observed in the replication study were consistent with those reported in Study 2 and are taken as confirmation that the differences in quiz performance observed in Study 2 between the two quiz groups could be attributed to a difference in quiz type rather than subject ability.

## CHAPTER 5

### DISCUSSION

A comprehension mismatch was defined in the present investigation as a mismatch between the learner's encoding processes while studying and the retrieval processes required to perform a criterial task. Since previous research has reported that college-age subjects will "orient" themselves to text material "at different levels" (e.g., factual/applied; dualistic/relativistic), this study was designed to examine learning outcomes under conditions in which subjects' comprehension goals did or did not match those represented by the criterion task. It was predicted that subjects would experience the greatest success under "match" conditions and that as comprehension goals diverged, performance would decrease. Other characteristics including reading proficiency, degree of dualism or relativism, interest in the article, and general mismatch experience were expected to serve as additional predictor variables in the analysis of quiz performance.

#### Comprehension Goals

Comprehension goals were assessed with a scale constructed of eight apparently factual goal items and eight apparently applied goal items. Factor analysis of this scale suggested, however, that the factual-applied comprehension goal measure was more appropriately described as a factual-main idea comprehension goal measure. Examination of the items comprising each of the primary factors led to the

conclusion that the comprehension goals seemed to be associated more on the basis of content, than on the type of cognitive processing presumed to influence the learning of that content. Using this interpretation it was observed that the majority of subjects obtained main idea comprehension goal scores in the medium to high range of the scale, while at the same time scoring at the low end of the scale for the factual comprehension goal orientation. Very few subjects could be classified into one individual category; technically speaking, most subjects had mixed comprehension goal profiles. Given the reinterpretation of the applied-main idea items on the measure, this finding may not be particularly surprising. That is, research consistently reports that college students who are good readers (as most of these subjects were), read more flexibly and for more purposes than poor readers and are generally competent at identifying the main ideas in a text (Brown & Smiley, 1977, 1978; Goetz & Armbruster, 1980; Hare, 1981).

The originally planned match-mismatch analysis of groups was modified somewhat to take into account the comprehension goal structure identified in the factor analysis. Using median splits on factual and main idea scores, subjects were classified as high or low on each comprehension goal dimension; the resulting analyses of quiz performance yielded some significant results with regard to the total and half-quiz scores. Of interest were the differences between the four groups of subjects with regard to the correlations between the NDRT and comprehension goal scores on the two types of quizzes.

Anderson and Armbruster (1984) have suggested that when students are told "that reading to remember the information in the text [is] the primary task" most students may presume this to mean that almost any information in the text could be tested on. Thus, attention and effort may be spread equally over the passage, perhaps reflecting the observation that students in the present study reported using a variety of goals while reading. Since the explicitness of the studying instructions or testing criteria may affect how an individual studies, specific instructions were avoided in the present research to encourage students to use as typical or natural a studying approach as possible.

Additional questions are raised by the format used in the comprehension goal measure. Previous research has generally involved the collection of data on subjects' use of study strategies as opposed to an assessment of their underlying comprehension goals. While the former may represent a more straightforward or objective response on the part of subjects, descriptions of study behaviors will not necessarily provide useful data about one's reading goals or the level of one's processing activities. The measure used in the present study was an attempt to get at that information from subjects directly. However, given the similarity in the importance ratings for the applied and factual versions of each content area, it is not evident that subjects readily made distinctions between the two orientations. It seems possible that the wording of some of the statements (e.g., "pay attention to the behaviors..."; "think about the behaviors of...in comparison to...") were not as salient or meaningful to subjects as predicted.

The lack of an expected relationship between comprehension goals and quiz performance also raises questions concerning the validity of the comprehension goal measure. Although the comprehension goal items were constructed to directly correspond to information presented in the quiz, it appears that it may have been too simplistic to assume, without preliminary testing, that this measure, which seemed to have sufficient face validity, also had sufficient predictive validity. These results demonstrate the necessity of having adequate measures for assessing nonobservable types of cognitive processing.

The previous discussion might be said to presume that subjects accurately reported their comprehension goals and that the failure to achieve results came at the measure level rather than the subject level. Alternatively, it could be argued that in fact one difficulty with the present study was its reliance, in part, on verbal report data. A number of researchers, for example, have questioned the extent to which a reader can accurately describe "the workings of her own mind" within the confines of a novel task situation (Afflerbach & Johnston, 1984; Cavanaugh & Perlmutter, 1982; Garner, 1982; Phifer & Glover, 1982). The characteristics of the reporting task--whether subjects are asked to use retrospective or concurrent reporting, as well as the use of probes or open ended questioning--will have considerable influence on the types of reading activities reported.

Although retrospective reporting as used in the present research is less disruptive to the reading task, students' reports of their strategy use after reading may be distorted by their perceptions of the researcher's goals, perceived task demands, or the subject's ability to

recall what he or she was thinking about while reading. This potential disadvantage may be compounded by the use of probing or a structured questionnaire. When subjects are cued to recall their use of specific comprehension goals as opposed to responding to an open ended question instructing them to describe what they did, there is concern that subjects will provide information about what they ought to do rather than about what they did (Garner, 1982). Cueing to particular strategies may be perceived by subjects as positive valuing and influence them to favor the reporting of certain strategies over others. This may have been especially likely in the present study since some of the comprehension goal items represented main ideas from the passage. In retrospect, it might have been more useful to probe subjects for simple behavioral descriptions of their activities using nonspecific language, despite the increased complexity of scoring.

Although certain disadvantages have been identified in the use of verbal reports in reading research, a number of investigators have argued that its particular advantages--including the opportunity to study data that would be difficult to obtain under other research conditions--supports its use (Afflerbach & Johnston, 1984). It is suggested here, however, that the nature of the verbal reporting task must be carefully considered. It is probably more difficult for a subject to describe how they studied particular parts of a text than to identify the behaviors they engaged in while studying. The use of a structured questionnaire will in addition create the possibility that certain activities the subject engaged in while reading will not be measured; this appears to be especially likely given the diversity of



study monitoring activities.

#### Quiz Condition Differences

Unfortunately it cannot be known whether any of these concerns were influential with regard to subjects' verbal reports. It is possible that subjects in the present research did not necessarily employ the techniques they claimed to have used. Alternatively, the comprehension goal measure, as constructed, may not have validly represented the criteria necessary to perform the tasks successfully. A third possibility appeared with the observation that in the case of the factual quiz, six of the comprehension goals entered into the regression equation individually did contribute in the prediction of quiz performance. This finding suggests that the comprehension goal measure may have been more appropriate for the assessment of factual quiz performance than applied quiz performance. The six comprehension goals included the factual and applied versions of the statements on child abuse, the statements on infant competency, and the statements on parental behavior. Together these six items accounted for approximately 31% of the variance. Why these particular comprehension goals were associated with quiz performance and others were not is difficult to determine. The statements on infant competency and parental behavior reflected two of the major concepts expressed in the article. The third topic--child abuse--received relatively low mean importance ratings overall but had proportionately higher standard deviations than most of the other comprehension goals.

It is possible that ratings of these statements somehow differentiated those students who read the passage carefully (and therefore thought about the important implications for child abuse) from those who did not. In contrast, the lack of a relationship between the comprehension goals and scores on the applied quiz suggests that a different set of operations underlay successful performance on that criterion task. Possibly, the task was so difficult that despite subjects' studying efforts, success was limited. It is also possible that subjects' studying efforts, though appropriate for the difficulty level or the nature of the comprehension goals on the factual quiz, did not match with the retrieval requirements of the applied quiz.

#### Dualism and Relativism

The dualism-relativism scale was intended to be used as an alternate measure of comprehension goal orientation using a nontext-specific format. Previous research has shown a relativistic orientation to be associated with use of applied as opposed to factual learning strategies and to be linked with greater academic success in college (Ryan, 1984). In the present research it was observed that scores on the Adherence scale showed a moderate negative correlation with reading ability such that higher reading scores were associated with the relativistic end of the continuum. Contrary to Ryan's (1984) observation that relativistic students are more successful on academic tasks, the present data indicated no such finding when controlling for subjects' reading abilities. It was furthermore observed that subjects classified as relativists obtained significantly higher reading scores than subjects classified as dualists, and that females showed a greater

tendency toward relativism than males. With regard to Ryan's (1984) findings, it appears possible that the relationship he observed between relativism and academic success might have been as readily explained in terms of subjects' reading proficiency. Reading ability has been consistently linked with academic performance (Zimmer et al., 1979).

The appearance of gender differences in the dualism-relativism scores does not appear to have been documented previously in the literature. Although females have frequently been shown to have greater verbal ability than males, the notion that the two sexes differ with regard to learning styles or cognitive processing has not been consistently established (Maccoby & Jacklin, 1974). In her recent work on women's moral development, Gilligan (1982) has argued that women perceive their social reality in a fashion that is qualitatively different from men, and that although men and women both experience a change in perspective from dualism to relativism in early adulthood, they reach the relativistic end of the continuum along different developmental paths.

It is possible that women in the present sample began their transition to relativism earlier than men and that the differences observed here are not apparent in later years. In describing her observations of the changes in moral understanding that occur in both men and women during the years following college, Gilligan (1982) notes that both sexes move away from absolutes during this time; however, the nature of the absolutes themselves differ for men and women. Perhaps the earlier growth toward relativism for college women is explained in part by their tendency to wrestle with intimacy needs prior to or

simultaneously with vocational choices rather than the typically observed "male" pattern of resolving identity issues prior to intimacy concerns (Hodgson & Fischer, 1979). Considering that college itself represents a particular vocational orientation, women who are also attempting to make compatible decisions concerning relationships, family and career may learn that existing dualistic constructs are insufficient for handling these new challenges and may accommodate to them by altering their cognitive structure.

The relationship between reading proficiency and scores on the Adherence scale is initially surprising since it does not appear to have been discussed previously in the literature (Perry, 1981). Whereas reading proficiency is viewed as a relatively stable trait following adolescence, growth along the Adherence continuum is predicted to continue into early adulthood (Perry, 1970). The relationship between the two variables would not appear to be entirely unexpected given that good readers probably use a number of reading strategies that are relativistic in nature; that is, good readers would be expected to make logical inferences from the material they have read, analyze and compare ideas from various sources and adjust readily and flexibly to a variety of reading-studying conditions.

#### Mismatch Experiences

In addition to examining the relationship between comprehension goals and testing experiences, an important goal of the present research was to gather additional data with which to better understand the comprehension mismatch experience. Although it had been predicted that subjects reporting frequent mismatch episodes would obtain lower reading

proficiency scores, this comparison was not significant at the .05 level. Nor did subjects varying in experienced mismatch frequency differ significantly with regard to Adherence scores, comprehension of the article, comprehension goals, or studying satisfaction.

Since the mismatch experience has been described as a discrepancy between "feelings of knowing" and performance on a subsequent criterion task, it was predicted that subjects reporting frequent mismatch experiences would be least successful at judging their own pre-quiz comprehension level. Consistent with this prediction it was observed that quiz performance was not correlated with pre-quiz comprehension for the high mismatch subjects, but was observed in the case of the moderate and low mismatch subjects. A subsequent stepwise regression, however, revealed that pre-quiz comprehension of the article proved to be a significant predictor of quiz performance only in the case of the moderate mismatch subjects. Generally, the regression analyses revealed different patterns of relationships for the three mismatch groups with regard to quiz performance. Reading proficiency, for example, proved to be a significant predictor variable in the case of moderate and high mismatch subjects; interest in the article was significant in the prediction of quiz scores for the low and frequent mismatch groups but not in the case of the moderate mismatch subjects. Although these patterns are interesting, they are difficult to interpret with any consistency. It is also important to emphasize that, due to sample size, these regressions were calculated using z-converted quiz scores and thus, the distinction in knowledge orientation between the factual and applied quizzes was lost. It does seem possible that these three groups

of subjects are different but it is not clear that the source of this difference--why they experience various levels of mismatching--has been identified.

#### Post-Quiz Evaluations and Mismatching

One significant, though unexpected finding in the present study, was the observation that subjects, irrespective of their comprehension goals or quiz performance, evaluated the applied quiz more negatively than the factual quiz. This finding was consistently observed in all of the post-quiz measures and seemed to be consistent with the observation that pre-quiz comprehension of the article was related to factual quiz performance but not applied quiz performance. Together these findings suggest that subjects' expectations concerning the type of criterion task they would receive were better met by the factual than the applied quiz. In actuality, the applied quiz appears to have created a mismatch experience for subjects insofar as these expectations were concerned. Technically, this mismatch cannot be described as a mismatch of comprehension goals since that analysis could not be performed. Examination of subjects' open ended descriptions of their quiz experiences indicates that while almost half of the subjects taking the factual quiz described it as "fair" or "appropriate," only a quarter of those taking the applied quiz described it similarly. One straightforward explanation for the higher ratings of the factual quiz might point to the higher scores on this quiz. However, it isn't clear given the lack of a relationship between quiz ratings and quiz scores, that this is a reasonable explanation. Although it seems likely that difficulty level did play a role in the quiz evaluations, the response

from students that "it just wasn't what I expected" implies that difficulty was not the only factor involved in the ratings. In part, there appeared to be a violation of subjects' expectations concerning the criterion task; it is possible, for example, that subjects in the present study operated with a less stringent or demanding definition of "comprehension" than they might have under other circumstances (Garner & Alexander, 1981).

#### Reader Interest

The importance of reader interest in text processing research was confirmed by the present findings which indicated that interest in the passage played a significant role in predicting performance on the applied quiz. It is possible that interest may be of particular importance in those cases where the passage and/or the criterion task are of a high-level or abstract nature. As the present findings indicate, reader interest may play a lesser role when the criterion task is less difficult or of a lower knowledge orientation. Olshavsky (1976-77), for example, reported a tendency for readers to utilize more monitoring strategies when they are interested than when they are not; thus, material that is interesting will be read more carefully, more incidental information is likely to be learned, and the content will be more readily recalled (Goetz & Armbruster, 1980). Interestingly, an informal survey of most contemporary text-processing research suggests that the use of reader interest as an independent variable is rare.

### Directions for Future Research

Returning to the title of this dissertation, "What is it that you know when 'you know'?" some preliminary conclusions would seem to be appropriate. First, given the relatively high prequiz comprehension ratings from subjects it seems safe to suggest that most of the students participating in the present research felt like they "knew the information." It is more difficult to identify what they knew. They performed well on the factual quiz and appeared to anticipate, in general, a criterion task that was comparable to that kind of quiz. More of the subjects in the factual quiz group, for example, felt that the quiz was fair and indicated that it could be said to assess the level of their comprehension reasonably well. For the applied subjects, all of the post-quiz evaluations were more negative than the factual quiz ratings; this was consistently observed in both Study 2 where subjects received only one quiz type and the replication study where subjects received all of the questions from both quizzes. These negative evaluations are attributed to the discrepancy between subjects' expectations of the task and the task itself.

Several important findings were revealed by the present data. First, it seems evident that comprehension is not a unitary construct. Additional research is needed to determine how various factors interact with the measurement of comprehension. These data point to the importance of reader interest, reading proficiency, subjects' perceived feelings of comprehension, and type of comprehension measure. The importance of the latter variable cannot be overestimated. It seems evident that significant findings with respect to one measure of



comprehension cannot necessarily be generalized to other operational definitions of comprehension.

With regard to continued investigation of the comprehension mismatch phenomena, it appears important to continue to explore this phenomenon with special attention given to its occurrence in natural settings as opposed to contrived or artificial situations. The findings presented here appear to point to the importance of the learners' expectations in creating a match or mismatch. It is predicted that these expectations will subsequently serve to guide the reader's encoding processes, which in turn are expected to influence performance on the criterion measure. Future research on the mismatch phenomenon might be best directed toward developing profiles of subjects differing in mismatch experiences with particular attention paid to learning styles, specific courses where mismatching seems to occur frequently or rarely, testing situations that appear to be correlated with mismatching, previous educational experience, and additional measures assessing subjects' abilities to judge the adequacy of their comprehension under a variety of learning and testing situations. It seems certain that the comprehension mismatch experience--or learning mismatches in general--are especially complicated; at the same time it is a problem of recognized importance in academic settings.

**LIST OF REFERENCES**

## LIST OF REFERENCES

- Afflerbach, P. & Johnston, P. (1984). Research methodology: On the use of verbal reports in reading research. Journal of Reading Behavior, 16, 307-322.
- Alvermann, D.E. & Ratekin, N.H. (1982). Metacognitive knowledge about reading proficiency: Its relation to study strategies and task demands. Journal of Reading Behavior, 14, 231-241.
- Anderson, T.H. (1980). Study strategies and testing adjunct aids. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 483-502), Hillsdale, NJ: Lawrence Erlbaum.
- Anderson, T.H. & Armbruster, B.B. (1984). Studying. In P.D. Pearson (Ed.), Handbook of Reading Research (pp. 657-679). New York: Longman.
- Anderson, R.C. & Pearson, P.D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P.D. Pearson (Ed.), Handbook of Reading Research (pp.255-291). New York: Longman.
- Andre, T. (1979). Does answering higher-level questions while reading facilitate productive learning? Review of Educational Research, 49, 280-318.
- Andre, M. & Anderson, T.H. (1978-79). The development and evaluation of a self-questioning study technique. Reading Research Quarterly, 14, 605-623.
- Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. Journal of Reading Behavior, 11, 365-374.
- Baker, L. & Anderson, R. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. Reading Research Quarterly, 17, 281-294.
- Baker, L. & Brown, A.L. (1984a). Cognitive monitoring in reading. In J. Flood (Ed.), Understanding reading comprehension (pp. 21-44). Newark: International Reading Assn.
- Baker, L. & Brown, A.L. (1984b). Metacognitive skills and reading. In P.D. Pearson (Ed.), Handbook of reading research (pp. 353-394). New York: Longman.
- Bloom, B.S., Engelhart, M.D., Furst, F.J., Hill, W.H. & Krathwohl, D.R. (1956). Taxonomy of educational objectives: Cognitive domain. New York: McKay.

- Bransford, J.D. (1979). Human cognition: Learning, understanding and remembering. Belmont, CA: Wadsworth.
- Bransford, J.D., Franks, J.J., Morris, C.D. & Stein, B.S. (1979). Some general constraints on learning and memory research. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory (pp. 331-354). Hillsdale, NJ: Erlbaum.
- Brown, A.L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology: Vol. 1. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, A.L. (1980). Metacognitive development and reading. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), Issues in reading comprehension (pp. 453-481). Hillsdale, NJ: Lawrence Erlbaum.
- Brown, A.L. & Smiley, S.S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. Child Development, 48, 1-8.
- Brown, A.L. & Smiley, S.S. (1978). The development of strategies for studying texts. Child Development, 49, 1076-1088.
- Butterweck, J.S. (1926). The problem of teaching high school pupils how to study. New York: Columbia University.
- Calfee, R.C. & Curley, R. (1984). Structures of prose in the content areas. In J. Flood (Ed.), Understanding reading comprehension (pp. 161-180). Newark: International Reading Assn.
- Cavanaugh, J.C. & Perlmutter, M. (1982). Metamemory: A critical examination. Child Development, 53, 11-28.
- Chase, W.G. & Simon, H.A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.
- Cioffi, G. (1986). Relationships among comprehension strategies reported by college students. Reading Research and Instruction, 25, 220-231.
- Cosgrove, J.M. & Patterson, C.J. (1977). Plans and the development of listener skills. Developmental Psychology, 13, 557-564.
- Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.
- Davey, B. & Lasasso, C. (1984). The interaction of reader and task factors in the assessment of reading comprehension. Journal of Experimental Education, 52, 199-206.
- Dewey, J. (1910). How we think. Boston: D.C. Heath.

- Erwin, T.D. & Millikin, J.L. (1980). The relationship of the Nelson-Denny Reading Test to the scholastic aptitude verbal tests. Measurement and Evaluation in Guidance, 13, 169-171.
- Flavell, J.H. (1981). Cognitive monitoring. In W.P. Dickson (Ed.), Children's oral communication skills (pp. 35-60). New York: Academic Press.
- Flavell, J.H., Friedrichs, A.G. & Hoyt, J.D. (1970). Developmental changes in memorization processes. Cognitive Psychology, 1, 324-340.
- Flavell, J.H. & Wellman, H.M. (1977). Metamemory. In R.V. Kail & J.W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, NJ: Lawrence Erlbaum.
- Fredericksen, C.H. (1975) Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 7, 139-166.
- Fry, E. (1968). A readability formula that saves time. Journal of Reading, 11, 513-516.
- Gall, M.D. (1970). The use of questions in teaching. Review of Educational Research, 40, 707-721.
- Garner, R. (1982). Verbal-report data on reading strategies. Journal of Reading Behaviors, 14, 159-167.
- Garner, R. & Alexander, P. (1981). Use of passage cues in answering of questions: Investigation of the effects of an explicit external criterion on adults' studying behavior. Journal of Reading Behavior, 13, 335-346.
- Garner, R. & Anderson, J. (1981-82). Monitoring-of-understanding research: Inquiry, directions, methodological dilemmas. The Journal of Experimental Education, 50, 70-76.
- Garner, R. & Alexander, P. (1982). Strategic processing of text: An investigation of the effects on adults' question-answering performance. Journal of Educational Research, 75, 144-148.
- Gerow, J.R. & Murphy, D.P. (1980). The validity of the NDRT as a predictor of performance in introductory psychology. Educational and Psychological Measurement, 40, 553-556.
- Gibson, E.J. & Levin, H. (1975). The psychology of reading. Cambridge, MA: The MIT Press.
- Gilligan, C. (1982). In a different voice. Cambridge, MA: Harvard University Press.
- Glucksberg, S., Krauss, R.M. & Weisberg, D. (1966). The development of referential communication skills. In F.D. Horowitz (Ed.), Review

of child development research, Vol. 4. Chicago: University of Chicago Press.

- Goetz, E.T. & Armbruster, B.B. (1980). Psychological correlates of text structure. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 201-220). Hillsdale, NJ: Lawrence Erlbaum.
- Goldberg, S. (1981). Premature birth: Consequences for the parent-infant relationship. In E.M. Hetherington & R.D. Parke (Eds.), Contemporary readings in child psychology (pp. 12-20) (2nd ed.). New York: McGraw-Hill.
- Hakstian, A.R. (1971). The effects of type of examination anticipated on test preparation and performance. Journal of Educational Research, 64, 319-324.
- Hare, V.C. (1981). Readers' problem identification and problem solving strategies for high- and low-knowledge articles. Journal of Reading Behavior, 13, 359-365.
- Hare, V.C. & Pulliam, C.A. (1980). College students' metacognitive awareness of reading behaviors. In M.L. Kamil & A.J. Moe (Eds.), Perspectives on reading research and instruction. 29th Yearbook of the National Reading Conf. Washington, D.C.
- Hayes, D.A. & Diehl, W. (1980). What research on prose comprehension suggests for college skills instruction. Journal of Reading, 23, 656-661.
- Hodgson, J.W. & Fischer, J.L. (1979). Sex differences in identity and intimacy development in college youth. Journal of Youth and Adolescence, 8, 37-50.
- Holt, J. (1964). How children fail. New York:Dell.
- Jenkins, J.J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory (pp. 429-446). Hillsdale, NJ: Lawrence Erlbaum.
- Kendall, J.R., Mason, J.M. & Hunter, W. (1980). Which comprehension? Artifacts in the measurement of reading comprehension. Journal of Educational Research, 73, 233-236.
- Kintsch, W. & Miller, J.R. (1984). Readability: A view from cognitive psychology. In J. Flood (Ed.), Understanding reading comprehension (pp. 220-232). Newark: International Reading Assn.
- Knefelkamp, L.L. Widick, C.C. & Stroad, B. (1976). Cognitive-developmental theory: A guide to counseling women. The Counseling Psychologist, 6, 15-19.

- Kohlberg, L. (1981). The philosophy of moral development. San Francisco: Harper & Row.
- Kraus, M. (1984). A preliminary investigation of students' explanations for mismatching. Unpublished manuscript.
- Kulhavy, R., Dyer, J.W., & Silver, L. (1975). The effects of notetaking and test expectancy on the learning of text material. Journal of Educational Research, 68, 363-365.
- Kunen, S., Cohen, R. & Solman, R. (1981). A levels-of-processing analysis of Bloom's Taxonomy. Journal of Educational Psychology, 73, 202-211.
- Maccoby, E.E. & Jacklin, C.N. (1974). The psychology of sex differences. Stanford: Stanford University Press.
- Markman, E.M. (1977). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 986-992.
- Markman, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 50, 643-655.
- Markman, E.M. (1981). Comprehension monitoring. In W.P. Dickson (Ed.), Children's oral communication skills (pp. 61-84). New York: Academic Press.
- Mayer, R.E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers on learning from science prose. Journal of Educational Psychology, 75, 40-49.
- Meyer, B.J.F. & Rice, G.E. (1984). The structure of text. In P.D. Pearson (Ed.), Handbook of reading research (pp. 319-351). New York: Longman.
- Morris, C.D., Bransford, J.D. & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 16, 519-533.
- Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H. (1975). Statistical package for the social sciences. New York: McGraw-Hill.
- Olshavsky, J.E. (1976-77). Reading as problem solving: An investigation of strategies. Reading Research Quarterly, 12, 654-674.
- Orlando, V.P. & Hayward, K.G. (1978). A comparison of the effectiveness of three study techniques for college students. In P.D. Pearson & J. Hansen (Eds.), Reading: Disciplined inquiry in process and practice. 27th Yearbook of the National Reading Conf. Clemson, SC.

- Ornell, C.P. (1979). The problem of analyzing understanding. Educational Research, 22, 32-38.
- Pearson, P.D. & Johnson, D.D. (1978). Teaching reading comprehension. New York: Holt, Rinehart & Winston.
- Perry, W.G. (1968). Patterns of development in thought and values of students in a liberal arts college: A validation of a scheme. Cambridge, MA: ERIC Document Reproduction Service No. ED 024 315.
- Perry, W.G. (1970). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart & Winston.
- Perry, W.G. (1981). Cognitive and ethical growth: The making of meaning. In A.W. Chickering (Ed.), The modern American college (pp. 76-116). San Francisco: Jossey-Bass.
- Phifer, S.J. & Glover, J.A. (1982). Don't take students' word for what they do while reading. Bulletin of the Psychonomic Society, 19, 194-196.
- Piaget, J. (1926). The language and thought of the child. London: Routledge & Kegan Paul.
- Postman, L. & Senders, V. (1946). Incidental learning and generality of set. Journal of Experimental Psychology, 36, 153-165.
- Rickards, J.P. & Friedman, F. (1978). The encoding versus the external storage hypothesis in notetaking. Contemporary Educational Psychology, 3, 136-143.
- Rosenblatt, L.M. (1978). The reader, the text, the poem. Carbondale: Southern Illinois University Press.
- Rubin, S. (1975). Professors, students and the syllabus. The Chronicle of Higher Education, 62, 56.
- Ryan, M.P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. Journal of Educational Psychology, 76, 248-258.
- Samuels, S.J. & Kamil, M.L. (1984). Models of the reading process. In P.D. Pearson (Ed.), Handbook of reading research (pp. 185-224). New York: Longman.
- Sax, G. & Collet, L.S. (1968). An empirical comparison of the effects of recall and multiple choice tests on student achievement. Journal of Educational Measurement, 5, 169-173.
- Schultz, C.B. & DiVesta, F.J. (1972). Effects of passage organization and note-taking on the selection of clustering strategies and on recall of textual materials. Journal of Educational Psychology, 63, 244-252.



- Short, E.J. & Ryan, E.B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. Journal of Educational Psychology, 76, 225-235.
- Smith, H.K. (1967). The responses of good and poor readers when asked to read for different purposes. Reading Research Quarterly, 3, 53-84.
- Smith, S. (1982). Learning strategies of mature college learners. Journal of Reading, 24, 5-12.
- Stanton, H.E. (1978). Helping students learn: The improvement of higher education. Washington, D.C.: University Press.
- Svensson, L. (1977). On qualitative differences in learning: Study skill and learning. British Journal of Educational Psychology, 47, 233-243.
- Thorndike, E.L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. Journal of Educational Psychology, 8, 323-332.
- Titchener, E.B. (1915). A beginner's psychology. New York: The MacMillan Co.
- Trachtenberg, D. (1974). Student tasks in text material: What cognitive skills do they tap? Peabody Journal of Education, 52, 54-57.
- Wagoner, S.A. (1983). Comprehension monitoring: What it is and what we know about it. Reading Research Quarterly, 18, 328-346.
- Zimmer, J.W., Glover, J.A., Ronning, R.R. & Petersen, C. (1979). On the utility of the Nelson-Denny Reading Test as a covariate in research on prose-processing. Perceptual and Motor Skills, 48, 641-642.

## TABLES

Table 1

Means and Standard Deviations for Comprehension Reading and Interest Measures, Mismatch Frequency, and Studying Behavior By Gender

Measure	Gender			
	Females (n = 32)		Males (n = 28)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Grade Point Average	3.46	.36	3.29	.37
Reader Interest	2.25	.62	2.68	.78
Prior Knowledge	1.36	.52	1.30	.52
Rated Comprehension	3.64	.69	3.47	.63
Rated Reading Ability	4.42	1.04	4.90	1.01
Mismatch Frequency	4.39	1.21	4.13	1.32
Read, Reread	3.18	1.59	3.13	1.56
Highlight	4.24	1.40	3.47	1.49
Notetaking	2.70	1.29	2.37	1.31
Outline	2.79	1.27	2.13	1.34
Self-Questioning	2.15	1.07	2.37	1.03
Summarization	2.58	1.16	3.40	1.09
Study Guide	3.70	1.37	3.57	1.41
Survey	2.82	1.44	3.27	1.40
Use Examples	2.70	1.38	3.17	1.47
Use Own Words	2.45	1.27	2.60	1.24

Note: Responses based on a partial sample of 60.

Table 2

Percentage of Responses for Each Comprehension Goal Level by Response Category

Response Category	Comprehension Goal Level		
	Knowledge	Comprehension	Application-Analysis
Information present in article	25% (76)	20% (61)	11% (32)
In article; not important for understanding	18% (55)	1% (2)	1% (2)
Main idea	24% (74)	41% (124)	31% (90)
Not in article; irrelevant	3% (10)	5% (15)	15% (41)
Somewhat important	5% (14)	1% (2)	2% (5)
Opinion-oriented; more than one answer	1% (2)	3% (10)	11% (32)
Confusing or vague	5% (14)	9% (27)	5% (14)
Fair question	10% (31)	12% (34)	9% (27)
Other/No response	9% (27)	8% (26)	15% (41)
TOTAL:	100%	100%	100%

Note: Numbers in parentheses are frequencies. Percentages are combined across the five questions in each comprehension goal level.

Table 3

Correlations Between the Comprehension Goal Ratings, Comprehension and Interest Ratings, Rated Reading Ability, and Mismatch Frequency

Variable	Comprehension Goal Level		
	Knowledge	Comprehension	Application- Analysis
Rated Comprehension	-.05	.05	-.10
Reader Interest	-.09	.17	.06
Rated Reading Ability	-.14	-.09	.03
Mismatch Frequency	.05	-.11	.01

Note: Responses based on a partial sample of 60.

Table 4

Means and Standard Deviations for Performance Measures by Mismatch Groups

Performance Measure	Mismatch Group			
	Low Mismatch (n = 47)		High Mismatch (n = 51)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rated Comprehension	3.55	.58	3.57	.73
Reader Interest	2.40	.68	2.51	.81
Prior Knowledge	1.21	.41	1.45	.58
Knowledge Ratings	4.74	1.31	4.73	1.41
Comprehension Ratings	5.12	1.53	5.08	1.50
Application-Analysis Ratings	4.36	1.75	4.43	1.69
Read, Reread	2.96	1.14	3.22	1.71
Highlight	3.60	1.48	3.90	1.45
Notetaking	2.38	1.39	2.41	1.20
Outline	2.34	1.36	2.31	1.21
Self-Questioning	2.23	1.03	2.10	1.06
Summarization	2.85	1.14	2.75	1.23
Study Guide	3.23	1.32	3.53	1.49
Survey	2.87	1.54	2.98	1.30
Use Examples	2.70	1.08	2.82	1.28
Use Own Words	2.53	1.35	2.25	1.15

Table 5

Intercorrelations Among Mismatch Ratings, Comprehension,  
Rated Reading Ability and Some Study Behaviors

Variable	1	2	3	4	5	6	7	8	9	10
1. Mismatch Frequency		-.26	-.53***	.14	.06	-.25*	-.18	.13	-.08	-.24*
2. Rated Comprehension			.24*	-.09	.02	-.05	.08	.01	.14	.06
3. Rated Reading Ability				-.25*	-.04	.26*	.07	-.13	.02	.40***
4. Highlight					.20	-.15	-.04	-.14	-.04	.08
5. Notetaking						.07	.04	.24*	.08	.22*
6. Self-Questioning							.28*	.08	.35***	.28*
7. Summarization								.32**	.57***	.20
8. Survey									.41***	.08
9. Use Examples										.43***
10. Use Own Words										

Note: \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 6

Percentage of Subjects Who Chose Each Reason for Mismatch

Reason for Mismatch	Percentage
Instructor didn't tell students what they needed to know	14% (14)
Student didn't know what was important to study	48% (47)
Student doesn't know how to study well	16% (16)
Student didn't spend enough time studying	46% (45)
Student experiences test anxiety	40% (39)
Student doesn't have the ability to do well	0% (0)
Student felt the test was unfair	27% (26)
Student performs poorly on certain kinds of test questions	21% (21)
Student knew the information but not in the right way	48% (47)

Note: Numbers in parentheses are frequencies. Percentages are based on the number of times each explanation was chosen. Subjects had the option to choose up to three reasons. N = 98.



Table 7

Intercorrelations Among Comprehension and Interest Ratings, Mismatch Frequency, Reading Proficiency and Comprehension Goals

Variable	1	2	3	4	5	6	7	8
1. Rated Comprehension		-.16*	.48***	-.02	.26**	.06	.15	.25**
2. Dualism			-.11	.06	-.35***	.03	.00	.03
3. Reader Interest				.13	.07	.18*	.28**	.46***
4. Mismatch Frequency					-.07	.04	.10	-.03
5. NDRT score						.11	.05	.09
6. Factual Comprehension Goals							.31***	.15*
7. Main Idea Comprehension Goal								-.02
8. Prior Knowledge								

Note:  $N = 140$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 8

Means and Standard Deviations for Performance Measures by Gender

Performance Measure	Gender			
	Females ( $n = 74$ )		Males ( $n = 66$ )	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rated Comprehension	5.61	.84	5.26	.84
Dualism Score	24.51	7.99	28.47	6.29
Reader Interest	4.88	1.58	4.03	1.30
Prior Knowledge	2.89	1.37	2.32	1.39
Mismatch Frequency	3.92	1.29	3.89	1.35
NDRT Score	98.74	23.26	90.47	23.40
Factual Comprehension Goals	19.59	4.54	18.59	4.81
Main Idea Comprehension Goals	32.05	4.76	31.23	4.69

Table 9

Factor Analyses of Comprehension Goal Measures

Scale Items	Factor Names		
	Dualism- Relativism	Main Idea Orientation	Factual Orientation
<b>Adherence Items</b>			
D1	.56	.03	-.02
D2	.49	-.10	-.04
D3	.48	-.01	.04
D4	.71	-.06	.01
D5	.56	.07	.04
D6	.61	-.04	.00
D7	.60	.07	-.06
<b>Factual Items</b>			
F1	.00	-.19	.42
F2	-.02	.37	.03
F3	-.12	.16	.39
F4	-.06	.24	.56
F5	.16	.13	.52
F6	.02	.69	.14
F7	.02	.61	.15
F8	.13	.09	.17
<b>Applied Items</b>			
Ap1	-.06	.45	.19
Ap2	-.02	.59	.04
Ap3	.01	.04	-.07
Ap4	-.12	.09	.21
Ap5	.14	.28	.53
Ap6	-.18	.25	.19
Ap7	-.12	.28	.30
Ap8	.08	.51	.03
Percent of Variance accounted for	38.8%	32.6%	17.1%
Eigenvalue	2.95	2.48	1.30

Note: Individual scale items are identified in Appendices A and B.

Table 10

Correlations Between Factual (FA) and Main Idea (MI) Comprehension Goals, Quiz Scores (Total, Hard and Easy Halves), and NDRT Score

A. Low Factual, Low Main Idea Group

	1	2	3	4	5	6
1. FA		.36*	-.04	-.12	.05	.03
2. MI	.42*		.05	.01	.07	-.15
3. Quiz Score	-.10	.01		.90***	.92***	.34*
4. Hard Half	.07	.00	.87***		.67***	.20
5. Easy Half	-.30	.02	.66***	.20		.41*
6. NDRT Score	-.15	-.24	.48*	.41*	.32	

Factual quiz group below diagonal ( $\underline{n} = 20$ ); Applied above diagonal ( $\underline{n} = 26$ )

B. High Factual, High Main Idea Group

	1	2	3	4	5	6
1. FA		.25	.44	.09	.63*	-.13
2. MI	.21		.22	-.36	.45	.25
3. Quiz Score	.12	-.07		.82**	.83***	.44
4. Hard Half	.10	-.05	.85***		.67***	.16
5. Easy Half	.08	-.06	.66***	.17		.56*
6. NDRT Score	-.02	-.37*	.42*	.41**	.22	

Factual quiz group below diagonal ( $\underline{n} = 30$ ); Applied above diagonal ( $\underline{n} = 10$ )

C. Low Factual, High Main Idea Group

	1	2	3	4	5	6
1. FA		.25	.29	.21	.25	-.03
2. MI	.02		.12	.04	.16	.06
3. Quiz Score	-.36	-.20		.86***	.62**	.81***
4. Hard Half	-.49	.23	.77**		.14	.59**
5. Easy Half	.18	-.63*	.37	-.32		.67**
6. NDRT Score	.20	.07	.00	-.01	.01	

Factual quiz group below diagonal ( $\underline{n} = 10$ ); Applied group above diagonal ( $\underline{n} = 16$ )

D. High Factual, Low Main Idea Group

	1	2	3	4	5	6
1. FA		-.30	.37	.48*	.12	-.08
2. MI	-.05		.26	.05	.26	-.04
3. Quiz Score	.08	-.18		.49*	.83***	.32
4. Hard Half	-.03	.06	.69***		-.08	.05
5. Easy Half	.13	-.31	.66***	-.09		.33
6. NDRT Score	-.18	.16	.13	.14	.02	

Factual quiz group below diagonal ( $\underline{n} = 15$ ); Applied group above diagonal ( $\underline{n} = 13$ )

Note: \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 11

Means and Standard Deviations for Performance Measures  
by Mismatch Group

Performance Measures	Mismatch Group					
	High ( <u>n</u> = 45)		Moderate ( <u>n</u> = 43)		Low ( <u>n</u> = 52)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Reader Interest	4.82	1.54	4.07	1.66	4.48	1.49
Rated Comprehensioncore	5.40	.83	5.42	.81	5.48	.89
NDRT Score	91.33	22.92	95.72	22.68	97.15	23.42
Main Idea Comprehension Goals	33.02	4.82	30.93	4.69	31.10	4.77
Factual Comprehension	19.66	4.44	18.72	4.72	19.10	4.56
Dualism	26.58	7.81	26.53	7.21	26.08	7.34

Table 12

Correlations Between Quiz Scores and Performance Measures for High, Moderate and Low Mismatch Groups

Performance Measures	Mismatch Group		
	High ( <u>n</u> = 45)	Moderate ( <u>n</u> = 43)	Low ( <u>n</u> = 52)
Rated Comprehension	.24	.39**	.36**
Reader Interest	.34**	.10	.47***
NDRT Score	.45**	.44**	.28*
Factual Comprehension Goals	-.08	.27*	.05
Main Idea Comprehension Goals	.04	.03	.15
Dualism	-.29*	-.12	-.30*
Reading Time	-.40**	.09	.11

Note: \* p < .05, \*\* p < .01, \*\*\* p < .01

Table 13

Stepwise Multiple Regressions for Low, Moderate, and High Mismatch Groups Predicting Z-Converted Quiz Scores

A. Low Mismatch ( $n = 52$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\overset{2}{\underline{R}}$	$\underline{df}$	$\underline{F}$
1	Reader Interest	.47	.98	.22	1,50	10.15***
2	Dualism	.54	.94	.30	2,49	10.28***
B. Moderate Mismatch ( $n = 42$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\overset{2}{\underline{R}}$	$\underline{df}$	$\underline{F}$
1	NDRT Score	.42	.82	.18	1,40	8.68**
2	Rated Comprehension	.52	.78	.27	2,39	7.34**
3	Reading Time	.60	.74	.36	3,38	7.02***
C. High Mismatch ( $n = 45$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\overset{2}{\underline{R}}$	$\underline{df}$	$\underline{F}$
1	NDRT Score	.44	.90	.19	1,43	10.13**
2	Positive Mismatch	.51	.86	.26	2,42	7.54**
3	Reader Interest	.59	.82	.35	3,41	7.25***

Note: \*  $p < .05$ , \*\*  $p < .01$ ,  $\underline{p} < .001$

Table 14

Stepwise Multiple Regressions for Prediction of Factual and Applied Quiz Scores

A. Factual Quiz Group ( $n = 75$ )						
Step	Variable Entered	<u>R</u>	<u>SE</u> EST	<u>R</u> <sup>2</sup>	<u>df</u>	<u>F</u>
1	Rated Comprehension	.36	1.69	.13	1,73	11.03**
2	NDRT Score	.45	1.64	.20	2,72	9.00***
3	F8	.52	1.57	.27	3,71	8.78***
4	F3	.58	1.51	.34	4,70	8.98***
5	AP5	.62	1.47	.38	5,69	8.58***
6	AP4	.66	1.41	.44	6,68	8.77***
7	AP2	.69	1.38	.47	7,67	8.56***
8	F6	.72	1.33	.52	8,66	8.86***

B. Applied Quiz Group ( $n = 65$ )						
Step	Variable Entered	<u>R</u>	<u>SE</u> EST	<u>R</u> <sup>2</sup>	<u>df</u>	<u>F</u>
1	NDRT Score	.44	1.92	.19	1,63	14.74***
2	Reader Interest	.54	1.82	.29	2,62	12.47***

Note: F8, F3, AP5, AP4, AP2 and F6 are identified in Appendix A.  
 \*  $p < .05$ ,  $p < .01$ ,  $p < .001$



Table 15

Means and Standard Deviations for Post-Quiz Measures  
by Quiz Condition

Measure	Quiz Condition			
	Factual ( <u>n</u> = 72)		Applied ( <u>n</u> = 63)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Post-Quiz Comprehension	5.14	.91	4.40	1.04
Pre- and Post-Quiz Comparison	3.69	1.03	4.62	1.22
Quiz Rating	5.22	.94	4.44	1.10
Question Ratings	61.78	6.23	56.09	5.89

Note: Five cases are missing from these analyses.

Table 16

Stepwise Multiple Regressions by Quiz Condition for Post-Quiz Evaluations

A. Factual Quiz ( $n = 72$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\underline{R}^2$	$\underline{df}$	$\underline{F}$
1	Post-Quiz Comp	.44	.83	.20	1,70	17.24***
2	Dualism	.55	.78	.30	2,69	14.88***
3	API	.59	.76	.35	3,68	12.23***

B. Applied Quiz ( $n = 63$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\underline{R}^2$	$\underline{df}$	$\underline{F}$
1	API	.38	.90	.14	1,61	10.00**
2	Positive Mis-match	.45	.87	.21	2,60	7.82**

C. Combined Sample ( $n = 45$ )						
Step	Variable Entered	$\underline{R}$	$\underline{SE}_{EST}$	$\underline{R}^2$	$\underline{df}$	$\underline{F}$
1	Post-Quiz Comp	.46	.91	.21	1,133	35.58***
2	Quiz Condition	.51	.88	.26	2,132	23.68***
3	Dualism	.54	.87	.29	3,131	17.93***
4	Question Ratings	.56	.85	.32	4,130	15.14***

Note: Five cases are missing from these analyses.

\*  $p < .05$ , \*\*  $p < .01$ ,  $\underline{p} < .001$

Table 17

Percentage of Responses by Category for Post-Quiz Evaluations

Rating Category	Quiz Condition	
	Factual ( <u>n</u> = 74)	Applied ( <u>n</u> = 64)
Too much emphasis on specifics and memorization	15% (11)	8% (5)
Some quiz items were confusing; had more than one answer	3% (2)	16% (10)
Some quiz items were opinion-oriented or required too much generali- zation	5% (4)	14% (9)
Quiz was too short to adequately assess under- standing	8% (6)	3% (2)
Multiple choice quizzes are not good measures	4% (3)	1% (1)
Quiz was appropriate	49% (36)	25% (16)
Quiz questions weren't what was expected	7% (5)	9% (27)
Other	9% (7)	12% (8)
TOTAL	100%	100%

Note: Two cases were missing from these analyses. Numbers in parentheses are frequencies.

## APPENDICES

APPENDIX A

Comprehension Goal Items

While reading the passage, how important for your understanding of what the article was about, was it for you to:

1. Focus on the names of researchers discussed in the passage (F1)
2. Pay attention to the differences in behavior between newborn preterms and newborn fullterms (F2)
3. Compare competence motivation and learned helplessness (AP1)
4. Learn the names of the three characteristics of the normally competent infant (F3)
5. Think about the behavior of parents when their preterm infant is a newborn in comparison to the behavior of parents when their preterm infant is older (AP2)
6. Imagine what it would be like to be the parent of a preterm infant (AP3)
7. Try to come up with some conclusions about the relationship between premature birth and child abuse (AP4)
8. Learn the definition of competence motivation (F4)
9. Learn the specific characteristics (e.g., weight), that determine whether an infant is preterm (F5)
10. Think about how the three characteristics of the normally competent infant differ (AP5)
11. Pay attention to the particular behaviors that parents of older preterm infants display (F6)
12. Think about how competence motivation might be related to situations other than those described in the passage (AP6)
13. Figure out how certain infant characteristics might

be related to the parents' feelings of effectiveness (AP7)

14. Pay attention to the particular behaviors that parents of newborn preterm infants display (F7)

15. Focus on the facts concerning preterm infants and child abuse (F8)

16. Determine what the author means when she describes "normal" infants as competent (AP8)

APPENDIX B

Adherence Scale

1. If professors would stick more to the facts and do less theorizing one could get more out of college (D1)
2. The best thing about science courses is that most problems have only one right answer (D2)
3. It is annoying to listen to a lecturer who cannot seem to make up his or her mind as to what s/he really believes (D3)
4. It's a waste of time to work on problems which have no possibility of coming out with a clear-cut and unambiguous answer (D4)
5. Educators should know by now which is the best method, lectures or small discussion groups (D5)
6. For most questions there is only one right answer once a person is able to get all the facts (D6)
7. A good teacher's job is to keep his or her student from wandering from the right track (D7)

## APPENDIX C

Summary Table for Gender (2) X Condition (2) Analyses of Variance: Study 1 (n=60)

<u>Variable</u>	<u>Gender</u> F	<u>p</u>	<u>Cond</u> F	<u>p</u>	<u>G X C</u> F	<u>p</u>
Interest	7.29	.009	.499	.483	.021	.885
Knowledge	.177	.676	.370	.546	.160	.691
Pre-comp	1.90	.173	10.20	.002	1.96	.167
Mismch	.793	.377	.041	.840	.076	.784
Read/RR	.021	.884	.086	.770	.177	.676
Highlight	4.47	.039	.022	.883	.065	.799
Notetaking	1.21	.276	1.39	.243	.000	.986
Outline	4.61	.036	1.37	.246	.561	.457
Summarize	10.61	.002	.423	.518	.193	.662
Self-quest	.763	.386	.128	.722	2.03	.160
Survey	1.43	.237	.087	.769	.037	.847
Examples	2.62	.111	.004	.951	.670	.416
Own words	.134	.716	2.06	.156	.121	.729

Summary Table for One-way Analyses of Variance on Mismatch Frequency (2): Study 1 (N=98)

<u>Variable</u>	<u>F ratio</u>	<u>p</u>
Pre-quiz comp.	.013	.909
Reader Interest	.484	.488
Prior Knowledge	5.44	.022
Read/reread	1.22	.273
Highlight	1.07	.304
Notetaking	.012	.913
Outline	.011	.918
Summarize	.213	.645
Self-questioning	.414	.523
Survey	.141	.708
Examples	.256	.614
Own words	1.21	.275



Summary Table for Gender (2) by Condition (2) Analyses of Variance: Study 2 (N=140)

<u>Variable</u>	<u>Gender</u> F	<u>p</u>	<u>Cond</u> F	<u>p</u>	<u>G X C</u>	<u>p</u>
Pre-comp.	.306	.581	.003	.958	4.659	.033
Facts	3.51	.060	1.16	.283	.190	.660
Main Ideas	1.48	.226	1.16	.283	.020	.880
Adherence	7.27	.008	.377	.540	.047	.829
Interest	9.19	.003	2.79	.097	2.095	.150
Knowledge	6.42	.010	4.43	.037	4.071	.046
Mismatch	.122	.727	1.53	.218	.541	.463
NDRT score	2.98	.086	3.07	.082	.159	.690
Reading Time	0.00	.991	2.57	.111	.000	.996
Quiz Score	.696	.406	59.60	.000	.022	.882

with NDRT score and reader interest as covariates  
 with NDRT score and pre-quiz comp. as covariates  
 with NDRT score as a covariate

Summary Table for Mismatch (3) by Condition (2) Analyses of Variance: Study 2 (N=140)

<u>Variable</u>	<u>Mismatch</u> F	<u>p</u>	<u>Cond</u> F	<u>p</u>	<u>M X C</u>	<u>p</u>
Pre-comp.	.110	.896	.005	.945	.870	.421
Facts	.228	.796	1.23	.295	.502	.606
Main Ideas	2.47	.089	.275	.601	.564	.571
Adherence	.045	.956	.287	.593	.565	.570
Interest	2.58	.079	.661	.418	1.25	.290
Knowledge	.810	.370	.574	.450	2.43	.122
NDRT score	.436	.648	3.70	.057	.401	.671
Reading Time	1.25	.291	2.05	.154	3.47	.034
Quiz Score	.640	.529	43.18	.000	.248	.781