

PREDICTIVE ACCURACY OF THIRD AND FIFTH GRADERS
ON LINGUISTIC AND NON-LINGUISTIC COMPREHENSION MEASURES:
THE RELATIONSHIP BETWEEN PERFORMANCE ON GRADUATED
PASSAGE QUESTIONS UNDER LISTENING AND READING CONDITIONS
AND THE STANDARD PROGRESSIVE MATRICES

by

Patricia A. Alexander

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

1981

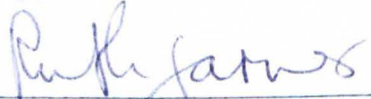
Copy 1

Approval Sheet

Title of Dissertation: PREDICTIVE ACCURACY OF THIRD AND FIFTH GRADERS
ON LINGUISTIC AND NON-LINGUISTIC COMPREHENSION
MEASURES: THE RELATIONSHIP BETWEEN PERFORMANCE
ON GRADUATED PASSAGE QUESTIONS UNDER LISTENING
AND READING CONDITIONS AND THE STANDARD
PROGRESSIVE MATRICES

Name of Candidate: Patricia A. Alexander
Doctor of Philosophy, 1981

Dissertation and Abstract Approved: _____


Dr. Ruth Garner
Assistant Professor of Education
Department of Early Childhood-
Elementary and Secondary Education

Date Approved: 9-24-81

Abstract

Title of Dissertation: PREDICTIVE ACCURACY OF THIRD AND FIFTH GRADERS ON LINGUISTIC AND NON-LINGUISTIC COMPREHENSION MEASURES: THE RELATIONSHIP BETWEEN PERFORMANCE ON GRADUATED PASSAGE QUESTIONS UNDER LISTENING AND READING CONDITIONS AND THE STANDARD PROGRESSIVE MATRICES

Patricia A. Alexander, Doctor of Philosophy, 1981

Dissertation directed by: Dr. Ruth Garner
Assistant Professor of Education
Department of Early Childhood-Elementary
and Secondary Education

This study, as an investigation of predictive accuracy, examined the relationship of subjects' performance on a non-linguistic measure of comprehension, the Raven's Standard Progressive Matrices, to their performance on linguistic measures of comprehension, graduated passages and questions from the Barnell-Loft Specific Skills Drawing Conclusions Series. These linguistic passages and questions were presented under silent reading and listening conditions. All third graders and fifth graders at a suburban Maryland parochial school served as the study population from which twenty subjects at each grade level were randomly selected. This sample was comprised of eighteen males and twenty-two females. All subjects received all treatment levels which consisted of four tasks administered in two sessions. Session I included the administration of the Raven's, with the added verbalization of item C-9, and passages and questions presented in a silent reading mode. In Session II all subjects completed passages and questions presented in a listening mode and the Slosson Intelligence Test. The graduated passages and question were contained in Comprehension Inventory I and

II which were administered in counterbalanced format. Half of the subjects received Inventory I as the listening task and II as the silent reading task, while half received Inventory I as the silent reading task and II as the listening task. The first question under research in this study was whether there was a relationship between linguistic and non-linguistic predictive accuracy. A multiple correlation technique was used to ascertain the relationship between the Raven's and the listening and silent reading linguistic comprehension measures. A correlation coefficient significantly different from zero was achieved in the comparison of subjects' performance data for these tasks. The second question addressed by this study was whether the relationship between the non-linguistic and the listening linguistic measure of predictive accuracy would be stronger than the relationship between the non-linguistic and the reading linguistic measure of predictive accuracy. Through the use of a partial correlation technique it was determined that the relationship between the Raven's and the listening task was not statistically stronger than the relationship between the Raven's and the silent reading task. Both partial correlations, however, were determined to be significantly different from zero. The third question investigated by this study was whether there was a significant relationship between the Raven's, as a non-linguistic measure of predictive accuracy, and the Slosson, as a linguistic measure of intelligence. The coefficient produced by the Pearson Product-Moment technique was assessed to be significantly different from zero. Finally, the study sought to ascertain if there were developmental differences in predictive accuracy. A t test for

related samples produced significant differences for all outcome measures in the direction of the fifth-grade subjects. These findings confirmed the hypothesis that older students would perform the specified tasks at a higher level of predictive accuracy than younger students. Possible explanations of the outcomes of this investigation, and their relationship to research in the areas of language/thought, analogous reasoning, intelligence/cognition, reading/listening and comprehension/prediction were discussed, along with implications for theory, research, classroom and diagnostic practices.

DEDICATION

To my son, John Franklin, the light
of my life and to my parents who
have given me much more than life.

Acknowledgements

The completion of this study was dependent upon the assistance, advice and support of many people. I am particularly grateful to:

Dr. Ruth Garner, my dissertation director, who has assisted me in many ways as role model, teacher, advisor, collaborator, and friend.

Dr. John Madison, my advisor, who has given me the freedom to grow while providing me with his warm support and guidance.

Dr. Robert Wilson, for his interest, suggestions, encouragement and acceptance.

Dr. Joseph McCaleb, for listening and responding to my questions and needs with concern.

Dr. Douglas Coulson, for his expertise and theoretical inquisitiveness.

Thanks is also extended to:

Vicci Barrett, for all the sharing we have done over these two years and for her special friendship.

Janet Gold, Ron Reis, Shirley Wagoner who gave willingly of their time, assistance and support in the completion of this endeavor.

Joseph Czarnecki, Natalie Felsher, Jackie Haynes,
Roy Lafontaine, Marianne Pfeiffer, and Patricia
Russavage who understand and have shared this
experience with me.

Table of Contents

CHAPTER I INTRODUCTION AND RATIONALE	Page
Introduction	1
Rationale	5
Theory	5
Research	7
Classroom Practices: Instruction and Diagnosis	8
Summary	9
Definition of Terms	9
Research Questions	11
Summary	11
CHAPTER II REVIEW OF RELATED LITERATURE	
The Relation of Language and Thought	12
Summary	21
Intelligence and Cognition	23
Summary	39
Analogous Reasoning: Linguistic and Non-Linguistic.	40
Summary	50
Reading and Listening	51
Summary	60
Comprehension and Prediction	60
Summary	72
CHAPTER III METHODOLOGY	
Purpose	74
Subjects	74

Materials	.75
The Ravens Standard Progressive Matrices	.75
Graduated Passages and Inferential Questions	.77
The Slosson Intelligence Test for Children and Adults	.80
Construct Validity of Materials	.81
Procedure	.92
Session I	.94
Session II	.98
Scoring Procedures	100
Pilot Study - Procedure and Material	101
Statistical Analysis of Data	103
Language/Thought	103
Reading/Listening	104
Comprehension/Prediction	105
Summary	106

CHAPTER IV RESULTS

Quantitative Analysis of Data	107
Post-Hoc Data Analysis for Construct Validity	113
Qualitative Analysis of Data	114
Summary	116

CHAPTER V SUMMARY, DISCUSSION AND IMPLICATIONS

Summary	117
Limitations	119
Discussion	120
Language/Thought	121

Intelligence/Cognition	124
Analogous Reasoning	126
Reading/Listening	129
Comprehension/Prediction	131
Implications	133
Theory	133
Research	135
Instructional Practice and Diagnosis	137
Summary	139
REFERENCE NOTES	140
REFERENCES	141
APPENDICES	
Appendix A: Data from <u>Post Hoc</u> Pilot Study	148
Appendix B: Data from Procedural Pilot	152
Appendix C: Scoring Materials	156
Appendix D: Sample Linguistic and Non-Linguistic Items	162
Appendix E: Analysis of Data by Sex	169
Appendix F: Data from Study	171

List of Tables

Table	Page
1 Sample Scoring Sheet for <u>A Priori</u> Analysis of Prediction-Difficulty	87
2 Linguistic and Non-Linguistic Task Items and Prediction-Difficulty Level	90
3 Means and Standard Deviations of Outcome Measures	108
4 Correlations for Outcome Measures	109
5 Means and Standard Deviations of Outcome Measures by Grade Level	112
6 Means and Standard Deviations of Differences for High and Low Prediction-Difficulty Items	113

List of Figures

Figure		Page
1	Schematic Representation of the Prediction-Comprehension Cognition Interaction	6
2	Representation of Spearman's Two Factor Theory	28
3	Representations of the Multiple Factor Theory	30
4	Guilford's Structure of the Intellect Model	31
5	General Processing Model for <u>Raven's SPM</u> , as an Extended Non-Linguistic Analogy	82
6	General Processing Model for Reading/Listening Passage, as an Extended Linguistic Analogy	83
7	Criteria for Determining Degree of Prediction-Difficulty	86

CHAPTER I

INTRODUCTION AND RATIONALE

Human beings are constantly bombarded by information from the environment, and the survival and ultimate success of individuals are largely dependent upon their ability to react and to respond appropriately to that information. From time to time the responses of individuals may appear instinctive or random in nature. For the most part, however, persons' reactions result from an analysis of available information, an examination of existing response-alternatives, and a selection of those response-alternatives that seem more likely to lead to desired goals. It is this process of eliminating unlikely alternatives and of enhancing the probability of an appropriate response that is termed "prediction" (Smith, 1978). The pervasiveness of prediction is attested to in Smith's remark that, "Everyone predicts - including children - all the time" (p. 66). Yet, relative to its predominance and importance, far too little is known about the prediction process, though the last decade has shown some increase in the number of investigations in this area.

It would appear from the research, for example, that prediction plays a crucial role in the reading process (e.g., Miller & Kintsch, 1980; Smith, 1978; Trabasso, 1980). As an information processing task, reading requires the individual to receive, interpret and respond to information visually and linguistically presented. Consequently, as prediction is vital to the processing of information, in general, so it would seem equally vital to the processing of a more specific form of information - language.

Listening as a related task where linguistic information is communicated in spoken rather than written form would also seem facilitated by individuals' ability to make appropriate predictions (Wilkinson, 1980). Additional research on the relationship between reading and listening prediction is needed, however, to ascertain more fully the effect of presentation-mode on predictive accuracy. In general, individuals are called upon to respond appropriately to the information they read or hear, and their success at these tasks is heightened by their ability to select appropriately from among the various response-alternatives available to them. Therefore, the investigation of predictive accuracy is one means of examining the linguistic processes of reading and listening in more depth.

One common way to investigate individuals' predictive accuracy is to have them answer questions based on information they have read or heard. Questioning is a predominant method of assessment employed in the classroom setting (cf. Durkin, 1978) and a technique used by many educational researchers to assess comprehension (e.g., Andre, 1979; Trabasso, 1980). Justification for such procedures lies in the fact that analyses of the responses to questions, particularly those of an inferential type, allow educators and researchers to draw various conclusions about individuals' ability to recognize, recall, and/or infer from the information encountered. As it relates to the process of prediction, the presence of correct responses would seem to indicate that the readers/listeners had successfully eliminated unlikely response-alternatives and had, thus, predicted accurately.

However, those initiating a questioning strategy in the investigation of predictive accuracy may gain additional insights by

considering the process through which a particular alternative was selected, as well as the accuracy or inaccuracy of the product - the response. Snow (1974) has suggested that much valuable information may be lost to educational researchers who do not attend to "what goes on inside S's inside the experiment" (p. 279).

Understandably prediction, as hypothesis-testing or problem-solving, forms the "basis of comprehension" (Smith, 1978), and, as such, plays a major role in the readers'/listeners' ability to derive meaning from linguistic information. Yet, it should not be assumed that the relation between prediction and comprehension is language-bound. As strongly argued by Huttenlocher (1976), thinking while involving symbolic processes, "does not always involve the symbols of natural language" (p. 262). Natural language, the term by which she signifies the linguistic code itself, may be seen as only one means by which individuals' experiences can be symbolized. Further, "knowledge of the code is separate from knowledge of the world" (p. 264).

In accepting this theoretical position that language and thought are independent (cf. Huttenlocher, 1974, 1976; Macnamara, 1977) one can readily accept the notion that reasoning, anticipation, or hypothesis-testing - descriptors applied to the prediction process (Smith, 1978) - can occur in relation to non-linguistic information, as well. Consequently, the examination of response-selection in relation to non-linguistic information could contribute to a fuller understanding of the prediction process. Some of the recent investigations in non-verbal analogical reasoning have begun to explore the thought-processes involved in tasks incorporating non-linguistic coding systems (cf. Goldman & Bisanz, 1980; Sternberg & Rifkin, 1979). The results of

these studies speak to the issue of non-linguistic prediction and their findings warrant further investigation.

To predict accurately, individuals must be able to make sense of the information with which they are presented and to interpret that information based on the knowledge and skills they already possess (Bransford, 1979). This, in essence, is the interactive view of cognition, where learning arises as "a function of relationships between the materials to be learned and the learner's currently activated skills and knowledge" (p. 141). Variations in the interpretation and subsequent variations in predictions may ultimately arise from numerous factors internal or external to the individuals involved. Such factors that may produce variations in responses include age or educational level, sex, and intelligence of individuals, as well as mode of information-presentation. The consideration of these factors may furnish additional insights into predictive accuracy not presently available.

In general, it would appear that while recent research studies have contributed to some degree to knowledge of the prediction process several areas of investigation relevant to this process require further examination. First of all, it is important, if not necessary, for research into predictive accuracy to elaborate on the theoretical framework from which it arises - a framework structured around cognitive-processing theory. It is this theory that provides the foundation on which current theories of comprehension, learning and even intelligence are being formulated. Further, as an observable outcome of cognitive processing, predictive accuracy may function as a vital link between these seemingly unrelated constructs and allow for quantifiable analysis of their potential relationship.

Rationale

In view of the stated need to extend knowledge of prediction, this study was undertaken with objectives related to theory, research and instructional, diagnostic procedures.

Theory

Generally speaking, the examination of prediction has focused primarily on its relationship to reading comprehension. As noted, prediction is considered to be "the basis of comprehension" (Smith, 1978, p. 67). Earlier research on prediction in the reading process centered even more specifically on readers' ability to determine unknown words through the "simultaneous use of semantic, syntactic, and graphic cues" (Goodman, 1967). While some continue to investigate prediction at the word level (e.g., Allington, 1978, 1979; Pearson & Studdt, 1975) they have been joined by those who seek to examine prediction from a "larger-than-word" perspective (e.g., Fillmore, 1981; Miller & Kintsch, 1980; Smith, 1978; Trabasso, 1980). Whether at the word, sentence or concept level (Nelson, 1977) accurate predictions must be based, in part, on the appropriate "interpretation of experience" (Smith, 1978, p. 240) - what Smith identifies as comprehension.

However, to restrict the exploration of prediction to its relationship to reading comprehension is to ignore its corresponding relationship to the more global construct, cognition (cf. Bransford, 1979) under which both processes - comprehension and consequently prediction - are subsumed (see Figure 1). Further, it is through cognitive theory that the construct of intelligence can be more fully examined (cf. Resnick, 1976) - a construct of vast importance in

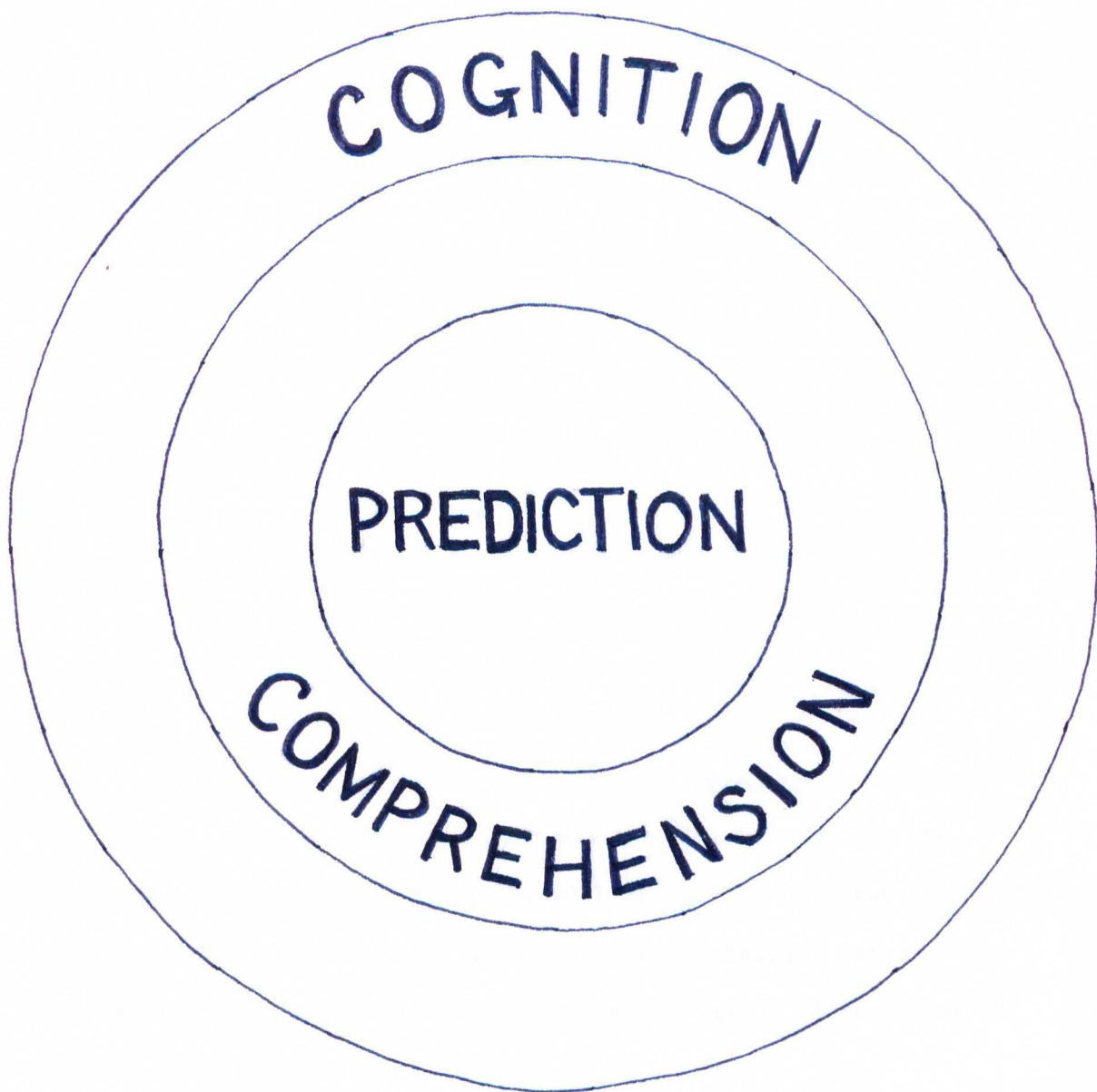


Figure 1

Schematic Representation of the Prediction -
Comprehension - Cognition Interaction

education today. It was the purpose of this study to take prediction beyond the bounds of reading comprehension and to explore more fully its relationship to intelligence and cognition. This global analysis of prediction seemed theoretically appealing in light of its potential relationship to these broader constructs of comprehension, intelligence and cognition.

Research

The present study was designed to contribute to the body of knowledge within the following areas of reading research:

Language/Thought - In this study predictive accuracy was assessed by subjects' performance across tasks identified as linguistic and non-linguistic comprehension measures. These measures were designed to require individuals to discern the relationship between the information presented and then to reason as to the best response to complete the tasks. Because the study furnished data relevant to individuals' ability to reason through both linguistic and non-linguistic information it contributed to the research on language and thought. The presence of a strong relationship between the language-based and non-language-based tasks was to be viewed as support for the theoretical position that language and thought can exist autonomously - a position expounded in the writings of Huttenlocher (1974, 1976) and Macnamara (1977).

Reading/Listening - The design of this study was such that all subjects performed parallel comprehension tasks presented under silent reading and listening conditions. A completion design

procedure (Wilkinson, 1980) was implemented. Supporting evidence was sought for those research studies that have found that listening-to-completion tasks produce better comprehension and memory than reading-to-completion tasks (Guthrie & Tyler, 1976; Rohwer & Harris, 1975; Rohwer & Matz, 1975). Comparison of performance on reading and listening comprehension tasks to performance on a non-linguistic comprehension task allowed for conclusions on the relative strength of the relationships to be drawn. In general, this study produced information regarding the accuracy of children's predictions relative to mode of presentation.

Comprehension/Prediction - The results of this investigation on predictive accuracy which examined the performance of third and fifth graders across various prediction measures also added to the current research on developmental differences in prediction. In addition, the study evaluated subjects' performance on inferential questions based on graduated passages. The specific passages incorporated in the reading and listening tasks can be considered "ecologically valid" (Snow, 1974), coming directly from instructional materials widely used in the classroom setting. This material organization contributed to knowledge on assessment through questioning strategies.

Classroom Practices: Instruction and Diagnosis

The analysis of predictive accuracy across linguistic and non-linguistic measures presented under silent reading and listening conditions provided direction for instructional, and diagnostic strategies. These implications for classroom practice will, it is hoped, eventually lead to a better understanding of the prediction

process and to learners' selection of more appropriate response-alternatives.

Summary

Rationale for the present study touched on three important areas - theory, research, and classroom practices.

Definition of Terms

The following terms are defined according to their use in the present study:

Prediction - The prior elimination of unlikely alternatives (Smith, 1978, p. 242), and the establishment of expectations based on the information discerned.

Predictive accuracy - The match in expectations between the sender of information (i.e., the generator of a particular instrument) and the receiver of that information (i.e., the subject in the investigation).

Cognition - A particular organization of knowledge in the brain, or the process of organizing such knowledge (Smith, p. 239).

Comprehension - The interpretation of experience; relating new information to what is already known (Smith, p. 240).

Non-linguistic comprehension - Understanding relatively uninfluenced by language symbols as evidenced by subjects' performance on the Raven's Standard Progressive Matrices, (Also referred to in this document as the Raven's and the SPM).

Linguistic comprehension - Understanding strongly influenced by language symbols as evidenced by subjects' performance on graduated passages from the Barnell-Loft's Drawing Conclusions (B-G) Series.

Information - Any property of the physical environment that reduces uncertainty, eliminating or reducing the probability of alternatives among which the perceiver must decide (Smith, p. 241).

Intelligence - The ability to learn even in the absence of formal instruction (Resnick & Glaser, 1980, p. 205), as measured by subjects' performance on the Raven's Standard Progressive Matrices and the Slosson Intelligence Test for Children and Adults, (Also referred to as the Slosson and the SIT).

Question - A direction to a learner to examine instructional material or his [her] memory of it and to produce some response (Andre, 1979, p. 281).

Inferential question - These questions are answered using the information found in the text, in addition to data previously observed and learned (Bush & Huebner, 1979).

Reading - Interaction of language user with textual information presented in a written mode (i.e., processing information silently read by the language user).

Listening - Interaction of language user with textual information presented in an aural mode (i.e., processing information orally read to the language user).

Research Questions

The following questions associated with the present study were the focus of this investigation:

Language/Thought

- Is there a relationship between linguistic and non-linguistic predictive accuracy?
- Is there a relationship between a non-linguistic measure of predictive accuracy and a linguistic measure of intelligence?

Reading/Listening

- Is there a stronger relationship between the listening linguistic predictive accuracy and non-linguistic predictive accuracy than between the reading linguistic predictive accuracy and non-linguistic predictive accuracy?

Comprehension/Prediction

- Are there developmental differences in linguistic and non-linguistic predictive accuracy?

Summary

The need for the study and the importance of each objective of the investigation have been presented. Definitions of terms and research questions indicate specific considerations and concerns of this investigator.

CHAPTER II

REVIEW OF RELATED LITERATURE

The account of related literature included here speaks to each of the three areas of concern in the current investigation - Language and Thought, Reading and Listening, and Comprehension and Prediction. The broad theoretical topic of Language and Thought has been divided into three more specific areas of inquiry - The relation of language and thought, intelligence and cognition, and analogous reasoning: linguistic and non-linguistic.

The Relation of Language and Thought

The relation of language and thought has been an issue of concern to investigators from such seemingly diverse disciplines as philosophy, psychology, anthropology and linguistics. The position on this issue taken by several noted researchers in the area of language and thought, among them Huttenlocher (1974, 1976), Macnamara (1977), Olson (1970) and Smith (1973, 1978), and the position upon which this study is based is that language and thought are autonomous systems. This is not to deny the fact that language and thought experiences often overlap, but it is a clear denial of the fact that language and thought are related in a one-to-one correspondence. This view of language and thought taken in this investigation is one which clearly differs from the theoretical

position of others such as Vygotsky (1962) and Chomsky (1957). To Vygotsky there was an unitary nature between thought and language, "a living union of sound and meaning" (1967, p. 4).

One of the first individuals to counter the long held claims of language and meaning unity, and to move away from a study of language focused in its syntax or structure was Chafe (1970; 1977). The message Chafe delivered and one which is still being expounded and investigated is that semantics or meaning is the "crucial component of language" (1970, p. 73). For a linguist to describe in detail the surface structure of a language without knowledge of the semantic structure was as Chafe (1970) said, like "trying to describe the production of milk while omitting the cow" (p. 73). Further, Chafe observed that there is no bi-directionality in relation to language and thought. One cannot progress from language to meaning in the same way that one can move from meaning to language. Language, according to Chafe, is but the symbolization of meaning constrained by the conventions of the medium.

Over a decade ago, Olson (1970) in a discussion of this very issue systematically distinguished between language and thought, or as he specified it - between an utterance and meaning. Olson arguing from a foundation in semantic theory stated that an utterance, in and of itself, conveys no meaning. Meaning, in actuality, is derived from the individual's knowledge of the perceived or intended referent of that utterance and is not determined by the syntactic or semantic markers that are part of the linguistic system. Meaning-getting, according to Olson, involves the language user's non-linguistic

knowledge. As support for the role of non-linguistic knowledge in semantic decisions, Olson presents several points of evidence.

First of all, employing the research of Chomsky (1957), Olson distinguishes between meaning and grammar. Meaningful sentences, it is demonstrated, can be non-grammatical in structure, and meaningless sentences, similarly, can be grammatically correct. Olson postulates from this, as had Chomsky, that meaning is distinct from grammar.

Additional support for Olson's contention comes from a consideration of pronoun and reflexive usage. It is shown that the choice between pronouns such as "he", "she", and "it" or "who" and "which" often depends on the intended referent of the noun phrase and not on the linguistic items in that phrase. Olson provides the following sentences to illustrate this point.

I bought the apple and I ate the apple.

I bought the apple and I ate it.

The man killed the man.

The man killed himself.

The provision of the second reflexive pronoun requires more than a simple recognition of the lexical entry, the man. The semantic decision must be based on the knowledge of the intended referent, as well.

The determination of ambiguity, anomaly and paraphrase are also dependent on the individual's knowledge of an intended referent. For example, it is impossible to specify, unambiguously, the meaning of a single word or sentence without presentation of context. This context

may be linguistic and/or non-linguistic in form. Likewise, in a paraphrase meaning of an utterance must be maintained while the language of the original utterance is altered to some extent. Finally, the anomaly of a message depends on the experience and imagination of the listener, where imagination is defined as "the ability to intellectually rearrange our perceptual experience" (p. 260). These three linguistic devices - ambiguity, paraphrase, and anomaly - appear to disavow any one-to-one relationship between an utterance and meaning.

Evidence for the role of extra-linguistic knowledge in semantic decisions also comes from the consideration of what Olson describes as ostensive definitions. An ostensive definition is that situation where an object is defined merely by pointing to the object in question. To illustrate, by pointing out "zebra" through a single word, the speaker represents all the characteristics of that object as well - "an animal resembling a horse but with a black and white striped body..." Certainly the thoughts conveyed may far exceed the single linguistic item uttered.

Finally, Olson observed that for the speaker, the experience or concept which might be specified in a linguistic code exists regardless of the verbalization act. The act of verbalization, in his view, does not furnish the speaker with any additional information. Language, in effect, does not form thought but transforms it.

Olson presents many convincing arguments for the autonomy of language and thought. However, his reasoning focuses on the utterance as the linguistic element of concern and the referent as the element

of meaning. This analysis, to adequately serve as the foundation for the theoretical distinction between language and thought, must be expanded to consider language and meaning from more global perspectives. Therefore, the relation of language and thought should also be examined within the reading as well as the listening act if it is to speak even more convincingly to the issue at hand.

Smith (1973, 1978) elaborated on several of Olson's arguments for the autonomy of language and thought, and carried these arguments into the realm of the reading process. Smith expanded on the notion of information by suggesting two categories - visual and non-visual - into which information may be divided. As a piece of print can serve as visual information in the reading process, so can an utterance function as visual information in the listening process. Basically, visual information is that part of communication that is available to the brain through the senses. Non-visual information, by contrast, is that knowledge already possessed by the reader/listener. This may be knowledge of the content of the message transmitted.

Smith also discusses two aspects of language that play a role in this consideration of the relation of language and thought. Smith refers to these two aspects as surface structure and deep structure. The surface structure, which is part of visual information, is comprised of the observable characteristics of the language, while the deep structure incorporates the meaning conveyed by that language. These two aspects are distinct components of the communication process according to Smith who states that "...there is not a one-to-one correspondence between the surface structure of language and meaning"

(1978, p. 71). Meaning lies beyond the marks on a page called language. It exists in the mind of the individual who interprets or transmits information by linguistic or non-linguistic means.

While the writings of Smith and Olson address the issue of language and thought well from a principally theoretical base they appear to furnish little quantifiable evidence to support their contentions. Bransford with his colleagues (e.g., Bransford, Barclay & Franks, 1972; Bransford & Franks, 1971; Bransford & Johnson, 1972), however, has examined the language and thought issue from a more experimental perspective. Throughout a series of experiments with Franks (1971) on the phenomenon of sentence acquisition and retention, Bransford has consistently observed that subjects acquired more general abstract information than what had been linguistically presented. This was an abstraction of linguistic information that, according to Bransford and Franks, demanded further investigation. Consequently, in a later series of experiments Bransford with Barclay and Franks (1972) and Johnson (1973) set out to explore "the interplay between linguistic inputs and extra-linguistic knowledge" (p. 383, 1972). The results of these studies seemingly support the hypothesis that one's ability to understand or interpret an utterance conveyed linguistically depends not only on what is seen or heard but on what the individual already knows or is capable of discerning from the extra-linguistic context.

In one experiment, for example, Bransford, Barclay and Franks (1972) describe subjects' tendency to encode more information than is actually present in the linguistic input. This increased knowledge

appears to have resulted from subjects' inferences about spatial relations. To illustrate, subjects considered the following sentence sets:

1. Three turtles rested beside a floating log and a fish swam beneath them.
2. Three turtles rested on a floating log and a fish swam beneath them.

After presentation of one of these two sentences, subjects were read recognition items and asked to indicate if these items had been heard during acquisition. They were also asked to rate their confidence in each response. The results on subjects' recognition of these two sentences indicated that the alteration of only two lexical items conveyed markedly different meanings to the listener. Those subjects receiving the second sentence were significantly more likely to falsely recognize that the fish swam beneath the log though no such linguistic input was given.

Johnson, Bransford and Solomon (1973) were also able to show that subjects recall more specialized information than that which was linguistically provided. One example of this is provided in the following sentences:

John was trying to fix the bird house.

He was pounding the nail when his father came out to watch him and to help him do the work.

Again, employing a false recognition paradigm, the researchers found

that subjects receiving that given linguistic input frequently recognized a hammer as being a part of the information heard. They apparently related the lexical element pounding to the instrument hammer - a link justified by non-linguistic experiences. However, the word hammer was not part of the actual linguistic input.

Finally, in testing subjects' recognition of related phrases, Bransford and Johnson (1973) used a similar false recognition paradigm to investigate subjects' justification of two events. Sentences similar to the following were used to provide the opportunity for justification.

The floor was dirty because Sally used the mop.

Critical recognition items were inserted to include the information subjects were likely to infer from experimental sentences similar to that given above. Subjects, it would appear from the results, elaborated on the linguistic input in order to make sense of the message with which they were presented. Therefore, meaning for the listener in this experiment was often an elaboration of the linguistic message received.

From these and additional experiments conducted by Bransford and his colleagues, it can be concluded that language is but the symbol system employed by individuals to convey a message. The thought from which that message arises or the meaning attached to that message may be an alteration, clarification or elaboration of the actual linguistic input. Consequently, an examination of how individuals create, comprehend, and remember linguistic utterances should go far beyond the linguistic system, itself, to the extra-

linguistic knowledge that mediates the message communicated. Generally what Bransford and his colleagues have demonstrated through their research is that the effective processing of language seems to require some pre-existing knowledge base. This prior knowledge may be derived from non-linguistic, as well as linguistic, domains.

However, Bransford's many experiments while informative on the issue of language and thought must be weighed against one principal limitation. The impressive results obtained by Bransford and his colleagues center on novel, creative materials. While such material permits the exploration of important questions in language and thought, it, itself, must be questioned in terms of its relevance to the type of linguistic information normally received and transmitted.

The research of Huttenlocher (1974, 1976) adds supporting evidence to the position of the autonomy of language and thought. Though Huttenlocher agrees that thinking involves symbolic processes, she strongly contends that the linguistic code - which involves sound patterns of lexical elements and the rules by which these elements are organized - is but one means by which experiences are symbolized. The linguistic code does bear a "systematic relation to conceptualization and thought" (1976, p. 262), but knowledge of the linguistic code is fundamentally different from conceptual knowledge - one's knowledge of the world. This is a position very similar to that of Bloom and Lahey (1978) who consider language to be an arbitrary though conventional code through which ideas about the world are represented in communicative contexts.

In general, Huttenlocher believes that information becomes part of one's experiences as a result of a mapping process where the

linguistic code is mapped onto non-linguistic perceptions of events. To be preserved information transmitted in a linguistic code must correlate with the non-linguistic experience of the information recipient. Therefore, "the study of comprehension cannot progress far without explicit recognition" that the ultimate meaning of the linguistic codes arise from the experiences of the language user (1974, p. 337).

Regarding symbolic systems, Huttenlocher also postulates that symbol schemes can be mapped directly onto the experiences or events which they represent without the intervention of natural language. Further, different symbol schemes, she states, seem to be used normally to encode information of different types. An individual may be able to employ one of those symbol schemes to represent information of one kind but not another.

Huttenlocher supports Bransford's conclusion that the comprehension of information must be explored in terms of individuals' conceptual knowledge and in terms of its larger linguistic and non-linguistic context. However, Huttenlocher provides a caution regarding those studies of comprehension where subjects are asked to describe simple schematic drawings involving a few easily named objects, properties and relations. Investigators must be careful not to rapidly conclude that the subjects' experiences with objects and events are always "encoded" in a manner that relates directly to natural language descriptions.

As a summation of the autonomous view of language and thought Macnamara (1977) describes thought as an abstraction with respect to the linguistic code. Knowledge of the world or conceptual

representations, therefore, need not "depend on any resemblance between themselves and the objects represented" (p. 2). Macnamara also outlines the attempts of three movements - artificial intelligence, generative semantics and case grammar - to produce complete mappings between the forms and structures of sentences and the abstract representations of thought. If this task were achieved, if there were a complete mapping between language and thought, then the current argument for the autonomy of language and thought would have little credibility. However, the form classes and organizational rules of language have not been - and perhaps cannot be - adequately specified in semantic terms. This can be taken as partial evidence for the distinction between language and thought.

Finally, Macnamara argues that the reason for the autonomy between language and thought is that thought takes place in the domain of "intentionality". Language, on the other hand, rests in the domain of conventionality and can only hint at the intended meaning. The relation of language and thought, according to Macnamara is analogous to the relation of evidence to conclusion. Perhaps Macnamara summarized the distinction between language and thought best when he stated that:

A child whose cognitive powers grew, but who failed to develop language would be a mute; one who developed language but did not express thoughts in it would be no better than a parrot (p. xi).

Summary

Language and thought are two distinct entities. Though individuals generally choose to communicate thoughts in a linguistic

code there is no one-to-one correlation between the two. It would appear from the evidence presented that one can "language" without thought. Similarly, one can think in the absence of language.

As there is no symmetrical relation between language and thought, there is no symmetrical relation between message and meaning. Meaning lies in the mind of the reader/listener. This message may be conveyed by a linguistic symbol scheme or by a symbol scheme which is non-linguistic. In interpreting a message the recipient of the communication may alter, clarify or elaborate the information received. This interpretation results from the mapping of the linguistic code onto the knowledge and experiences already possessed by the recipient.

Intelligence and Cognition

Intelligence according to Snow (1980b) is "now accepted as one of psychology's most central concerns", and "educational psychologists increasingly recognize that intelligence is education's most important product, as well as it's most important raw material" (p. 185). However, prior to the decade of the seventies, the investigation of intelligence was principally of interest to individuals in the discipline of differential or correlational psychology, and few others. Most educators had come to view intelligence, pragmatically, as whatever the score or level of performance on an intelligence test indicated it to be. Few questions were asked. As now, measures designated as intelligence tests were found to be consistent and basically accurate predictors of school success in America and Europe (Resnick, 1976).

In addition, these intelligence tests and "their close cousins, the various differential aptitude tests" (p. 2) could readily be administered and analyzed for large groups of individuals - making them especially attractive in light of mass education.

The pragmatic nature of intelligence tests and the corresponding identification of intelligence as the level of performance on such tests were placed into historical perspective by Tyler (1976). According to Tyler, the growing reliance on psychometric instruments, which reached its height in the 1920's, was a direct consequence of the compulsory education movement of the late 19th and early 20th centuries. Educators were suddenly overwhelmed by the need to educate masses of individuals and were ill-prepared to cope. Group-administered intelligence measures permitted the identification of individuals who would require "special" instruction (e.g., tracking). Mental-test developers such as Binet, who is credited with "fathering" the psychometric movement, were at least dealing with a real educational need.

Tyler (1976) also suggests that the pre-existing notion of class distinction made educators and the public more susceptible to the idea of intelligence as a test score or level of performance on psychometric measures. In essence, the IQ was a reinforcement to the flourishing concept of social order. The product became the critical concern, and there was little need to question or to look deeper into the process by which that score or level was obtained. "Score became all important" (Tyler, 1976, p. 16). This attention to product - not process - was tempered to a small extent by the clinical psychologists who focused more on the factors attributing to such scores, and who augmented the

intelligence tests with other clinical measures. Their influence was not enough to alter the trend, however. With the outbreak of world wars large-group screening devices were essential and there was neither time nor desire to consider the individual behind the score.

Yet, with the outset of the seventies there was evidence of a transition in the concept of intelligence (Resnick, 1976; Tyler, 1976). Among the factors contributing to this transition, Tyler (1976) listed the growing repertoire of sophisticated experimental procedures. Researchers in intelligence no longer needed to be bound to simple correlational studies. More complex tools were available to the researcher who wished to ask more complex questions. Resnick (1976) described the increasing skepticism as to what intelligence tests actually measure as an additional factor in this change. The practical and social utility of the intelligence test was apparent but the basis of its invention became far less clear. Further, while many psychometric tests were shown to correlate highly, there were individuals who came to question the nature of this correlation - a deeper issue than was being addressed by traditional psychometrics.

As perhaps one of the most influential factors bringing about the transition in the nature of intelligence, Resnick (1976) and Tyler (1976) singled out the expanding interest among experimental psychologists in the concept of individual differences - a concept they had abandoned for so many years to the differential psychologists. Coupled with this phenomenon was the appearance of an off-shoot of experimental psychology specifically concerned with the internal processes from which knowledge arises - cognitive psychology. It is cognitive psychology that has shifted the emphasis in intelligence from product to process.

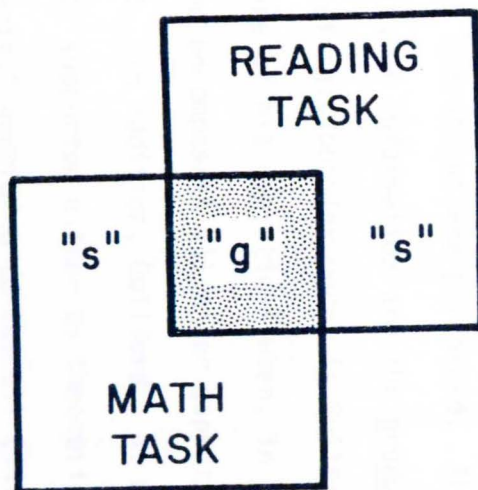
It should not be implied that cognitive psychologists have completely discounted the earlier labors of the psychometricians, though they have found these earlier approaches to intelligence to be product-focused. According to Carroll (1976) all of the key figures in psychometrics - Binet, Spearman, Thurstone and Guilford - demonstrated an abiding concern for the theoretical nature of intelligence and realized that "to construct a theory of intelligence is to construct a theory of cognition" (p. 27). Consequently, though the instruments devised by the psychometricians have been narrowly employed to identify intelligence, the theoretical framework they established helped structure the foundation for individual differences now being explored by cognitive psychologists. Further, there have been joint ventures between cognitive psychologists and psychometricians into the realm of individual differences in intelligence, and several studies reported here have come about as a result of this cooperative atmosphere (e.g., Carroll, 1976, 1980; Snow, 1980a).

Due to the relative importance of the early psychometric theories on the nature of the intellect and the role they played as precursors of more cognitive interpretations of intelligence, it is necessary to examine these theories briefly. The interest that these particular theorists - Spearman, Thurstone, Guilford, Cattell and Horn - demonstrated in the structure of the intellect later drew cognitive psychology into the consideration of individual differences in intelligence. The ideas their research generated still reflect in the labors of such current investigators into intelligence as Carroll (1976, 1980) and Sternberg (1977, 1980) whose studies link the

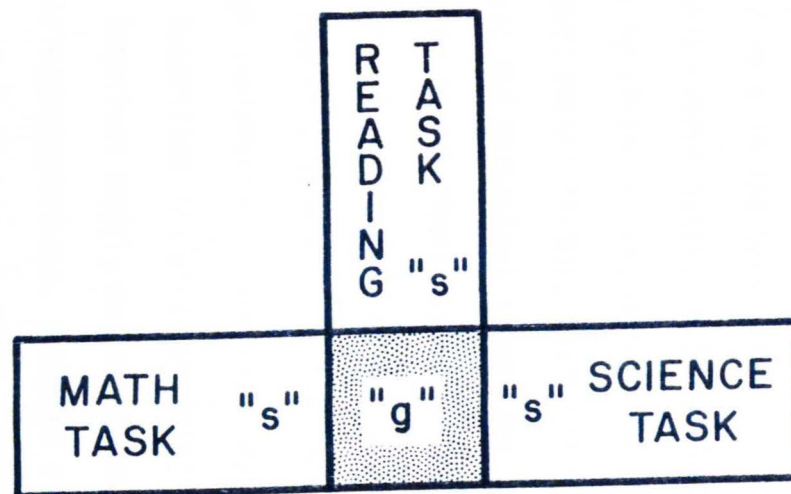
sophistication of experimental procedures with the sophistication of psychometric and cognitive theory.

One of the earliest attempts to discern the structure of intelligence was Spearman's Two Factor Structure Theory (1904, 1927), derived from results obtained from a factor analytic procedure developed by Spearman. In his theory, Spearman postulated that there were two types of intellectual aptitudes or factors - general and specific. General aptitude or the "g" factor of intelligence as Spearman referred to it was that aptitude common to all intellectual activities. Specific aptitudes or "s" factors, by comparison, were those factors specific to the completion of a particular task (see Figure 2). It was Spearman's contention that the "g" or general factor could best be measured through tasks of abstract reasoning such as the Raven's Standard Progressive Matrices. On the other hand, specific content-area aptitude tests were better at assessing the "s" factors of intelligence. Though the Two Factor Theory was a widely accepted interpretation of the intellectual structure repeated attempts by Spearman and his associates to identify the essence of the "g" factor in a more exact way never proved successful. Several cognitive psychologists - among them, Carroll (1980) - presently speculate that the "g" factor identified by Spearman may, in actuality, represent "a common pool of cognitive operations and strategies and of memory stores" (p. 60) that are required in the performance of most cognitive tasks.

Thurstone (1938) later expanded Spearman's concept of intelligence factors. In Thurstone's view there were many more aptitudes in the structure of the intellect than could be accounted for in a simplistic



A.



B.

Figure 2

Representations of Spearman's Two Factor Theory

two-factor theory. Intelligence, according to Thurstone, could be viewed as comprising a set of multiple factors or primary mental abilities that played a role in the performance of various tasks. The apparent relationships among tasks, in Thurstone's opinion, could be accounted for by seven primary mental abilities which included: number factor; verbal factor; space relations; memory; reasoning; work fluency; and perceptual speed (see Figure 3). Mental tests were constructed to assess these various primary abilities.

Guilford (1959) carried the concept of factor-analysis in intelligence to even greater distances. In his factor analytic model arising from a thorough exploration of the experimental psychological literature, Guilford identified 120 distinct intellectual aptitudes. This Structure of the Intellect model, one of the most prominent of the recent psychometric representations, presents a three way classification of these factors by operations, contents and products (see Figure 4). In Guilford's view operations are what organizers do in the processing of information, while contents represent the types of information on which operations are performed. The outcomes of the organisms processing of information are the products of the model. One of the weaknesses of this popular model is Guilford's classification of cognition as just one of his operations when, in actuality, cognition as currently understood encompasses all other aspects of the Structure of the Intellect model. In addition, Guilford's SI model does not seem "sufficiently well supported either by theoretical considerations or by the empirical facts," according to Carroll (1976, p. 29).

A final theory on intellectual structure that deserves consideration is the Crystallized-Fluid Model derived from the work of Cattell (1963,

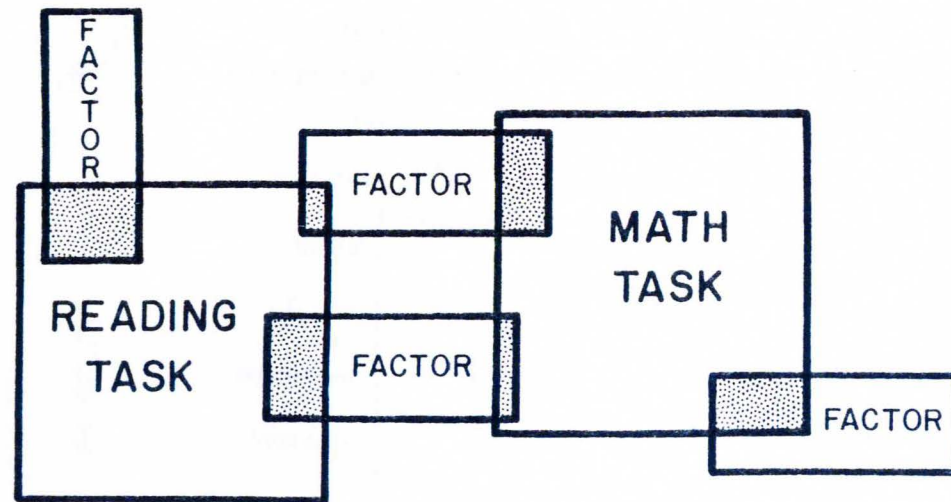


Figure 3

Representation of the Multiple Factor Theory

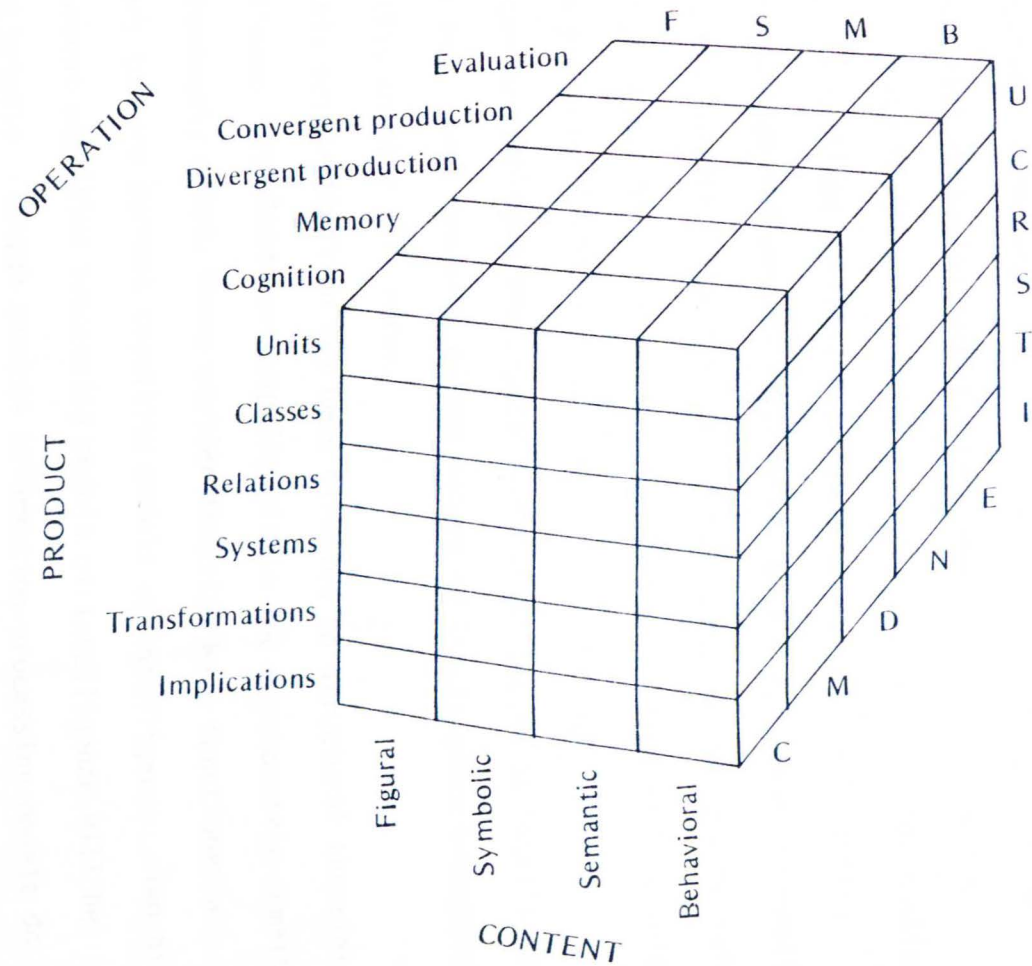


Figure 4
Guilford's Structure of the Intellect Model

1971) and Horn (1968). Attending even more to a cognitive theory base than other theoretical alternatives, the Crystallized-Fluid Model of the intellect identifies three parameters of intelligence: action phases, content and process parameters. Fluid ability, which is best assessed by tests of abstract reasoning such as the Raven's SPM relates to the individuals' ability to deal with current executions on novel information. By contrast, crystallized ability, best measured by vocabulary tests, reading comprehension and general information tests, taps that information of an accumulated nature. Fluid ability, therefore, can be said to refer more to the processes of the intellect, while crystallized ability refers more to the products already accumulated by the intellect. Snow (1979, 1980a) has employed the crystallized-fluid theory of intelligence in some of his recent research in the nature of intelligence - research to which this chapter will refer.

To this point attention has been directed to structural theories of intelligence postulated principally within the psychometric domain. As was previously noted, these earlier theories have functioned as precursors to more current cognitive models of intelligence. Two of these current cognitive processing models of intelligence will be outlined briefly. Though several information-processing models do exist in the literature the theories of Sternberg (1977, 1980) and Carroll (1976, 1980) will be emphasized because of their relative importance in the cognition literature and their direct relevance to the current study.

Carroll's new "Structure of the Intellect" model (1976), for example, is an obvious attempt to bring the practical value of psychometric tests into the theoretical framework of cognitive

psychology. His stated purpose is to demonstrate how various tasks contained in psychometric tests are cognitive tasks with identifiable structures, contents and central processes. To fulfill his purpose Carroll proposes a methodology for characterizing factor-analytic FA factors in relation to a cognitive processing model. Unlike the preceding factor-analytic models of intelligence Carroll's FA model begins with cognitive processing and derives the factors from that theoretical base - a methodology that is basically the reverse approach taken in most factor-analytic studies.

To construct his model Carroll went through a series of systematic procedures. First, a representative sample of 74 cognitive-psychometric tests measuring 24 FA factors was selected and a uniform system for coding the characteristics of the task was developed. Then, with the coding system programmed into an interactive computer a detailed analysis was made of 48 tests. Carroll considers the "essential results of this study" to be "the cognitive processes identified as being characteristic of each of the 24 FA factors and the tests that represent them" (p. 41). Carroll does admit that the characteristics of the FA factors are speculative but he believes that such speculation should be made as a result of factor analysis rooted in a theory-oriented investigation.

As compared to the previous intellect models, Carroll's FA model also employs factor analysis and would, upon that criterion, appear similar to those earlier approaches. However, Carroll's methodology is important in the investigation of intelligence for at least three reasons. First of all, the FA factor model begins with a particular theory base and allows the characteristics of the various factors that

arise to be inherent in that theory. As described, Carroll's model is rooted in cognitive, information-processing theory. Theory precedes factor characterization - Factor categorization does not precede theory. Secondly, Carroll's current research is an intentional link between the products of the psychometricians and the processes of the cognitive psychologists. In Carroll's view and in the opinion of others (e.g., Cronbach, 1957; Snow, 1980b) an understanding of intelligence can be best achieved through such cooperative efforts. Finally, the FA factor approach provides researchers and educators with a better sense of what commonly used tests are actually measuring. Through Carroll's methodology those employing psychometric tests have an increased awareness of the cognitive processes involved in successful completion of task items.

These various statements just made about the importance of Carroll's FA model are the reasons for its relevance to the present study. As with Carroll's methodology, the current study began its investigation of intelligence by moving outward from cognitive, information-processing theory. In addition, the study attempted to link cognitive theory results obtained from existing psychometric tests - the Raven's SPM and the Slosson Intelligence Test. Ultimately, it is the purpose of this study to investigate the processing tasks required in parallel linguistic and non-linguistic measures if subjects are to complete the tasks with a high degree of predictive accuracy - a purpose that corresponds to Carroll's work in the FA factor model.

Another variation of the factor-analytic approach employed by earlier psychometricians is now being developed and implemented by Sternberg (1977, 1980) in the examination of intelligence. However,

unlike Carroll, who altered the direction of investigation from factor to theory to theory to factor, Sternberg has also modified the unit of analysis. Rather than being concerned with factors of intelligence, defined as "the latent abilities of some kind that give rise to measured individual differences in performance on intelligence tests and in the real world" (1980, p. 6), Sternberg's research relies on the component. A component is identified as "an elementary information process that operates upon internal representations of objects or symbols" (p. 6). According to Sternberg, these components may be subdivided by function into: metacomponents, or higher order control processes, performance components, used in the execution of strategies for performance tasks, acquisition components, involved in learning new information, retention components, required in retrieving information and transfer components, involved in carrying information over from one situation to another.

Among various kinds of components, Sternberg also describes four interrelations. To begin with, one kind of component can activate another directly. Also, one type of component can activate another indirectly. Further, one component can directly or indirectly provide feedback to another. It is important to note that only metacomponents can activate or receive feedback directly from the other kinds of components.

In addition, cognitive processing components of the intellect theorized by Sternberg can be classified by their level of generality as well as by their function. By generality there are three types of components identified - general, class and specific. General components are processes that are used in completing all tasks within a particular task universe (algebra word problems constitute a task universe, for example). Class components are those processes used in the completion

of a proper subset of tasks within a task universe (multiplication of two-digit words problems, for example). Finally, specific components are processes used in the completion of a single task within a task universe.

The component model generated by Sternberg has relevance to the current investigation for two reasons. First of all, from a theoretical perspective, Sternberg's research, as with Carroll's, provides the foundation for the analysis of various cognitive processes required in accurate task completion. Secondly, in the analysis of performance components Sternberg has described a small set of components that are present in a wide array of inductive reasoning tasks. This class of performance components which include inferencing, mapping, encoding and application (prediction) are employed as indicators of the parallel nature of the linguistic and non-linguistic tasks incorporated in the current study. These various component processing terms will be discussed in more detail in this chapter as they relate to the study's specific methodology.

It was Sternberg's perception that componential analysis can provide educators with a diagnostic profile of the skills responsible for an observed deficiency. However, it would appear that predictive accuracy is perhaps better served by factor-analysis as applied within the more traditional psychometric model. Sternberg also observes that there is strong evidence that "componential training is possible, instructionally feasible, and educationally beneficial" (1980, p. 12).

In reaction to Sternberg's componential theory, Carroll (1980) remarked that like factor analysis, componential analysis is basically a method rather than a theory. By contrast, while factor analysis

requires the analysis of data obtained from a large number of subjects, componential analysis can be conducted on data obtained from one subject. Also, factors as source traits of individual differences are different from components which are essentially measures of processes. Finally, as Carroll's FA factor model made new use of existing psychometric information, Sternberg's component model of intelligence makes new use of factor-analytic data. Factors, as such, are not discarded but combined and considered within the various component processes defined by Sternberg.

Overall, the interpretation of intelligence within cognitive theory and the corresponding emphasis on intellectual processing has several obvious advantages over the preceding psychometric approach. Primarily, it brings together sophisticated experimental methodologies, as those employed by Carroll and Sternberg, with an abiding concern for the processes of the mind rather than for the products of the intellectual system. It seeks to look beyond the surface of psychometric tests to the internal operations and motivations from which the potential for success or failure arises. Yet, cognitive theory as applied to the nature of intelligence is not without its perceived limitations (Snow, 1980b). As Snow states, the study of intelligence has historically produced its monuments - measurement and factor theory, and mental test technology, for example - and information processing theory appears destined to become such a monument. What Snow suggests is a careful evaluation of the potential weaknesses of an intelligence-monument built upon cognitive theory. Among the possible weaknesses Snow lists the following:

1. The problem of static media - where the dynamic human intelligence is analyzed by means of static tasks.
2. The problem of representativeness - where intelligence is studied by means of tasks which are unlike real world tasks and which, consequently, produce unnatural behaviors.
3. The problem of collective intellect - where the investigation of human intellect ignores the collective intelligence that arises from interactions between individuals and concentrates only on individual intelligence.
4. The problem of conation - where there is no examination of the interaction of cognition and affect or the purposeful actions between individual and situation.
5. The problem of ambiguity and the need for playfulness - where the study of intelligence considers only problem-solving and not problem-finding, as well.

Some of these limitations outlined by Snow can be readily removed, in his opinion "by adapting standard psychometric methodology to experimental purposes" as has been done by Sternberg (1977, 1980) and Snow, himself, or by adapting "standard experimental methodology to psychometric purposes" (1980b, p. 186). By his speculation on the limitations of cognitive theory in the research on intelligence Snow does not wish to overlook the significance of that theory to the

increased understanding of intelligence. His stated purpose is simply to forestall "a premature molding of the next monument" (p. 186).

Summary

Intelligence is an evolving concept. From the work of Binet until the decade of the seventies intelligence has been understood primarily as a score or performance on a mental test. Since the seventies, however, intelligence has come to be interpreted, more and more, as effective and efficient cognitive processing operating interactively with an individual's prior experiences. In addition, it would appear that this move from a product-oriented to a more process-oriented view of intelligence has been spawned by such factors as an expanding repertoire of experimental techniques, an increasing interest in individual differences among experimental psychologists, and principally, by a growth of the branch of experimental psychology known as cognitive psychology.

As Snow (1980b) posited, cognitive psychology seems destined to become the next monument built in the exploration of intelligence. Snow cautions, however, that the construction of that monument must be carefully scrutinized for potential weaknesses. These weaknesses, he assumes, can structurally be reinforced through the cooperative labors of psychometrics and cognitive psychology. The knowledge contributed by both of these disciplines could possibly lead to a greater understanding of the nature of intelligence.

Analogous Reasoning: Linguistic and Non-Linguistic

Over fifty years ago, Thorndike (1917) concluded that reading is reasoning. This interpretation of reading as a reasoning process has recently been echoed by various researchers within cognitive psychology (e.g., Rumelhart & Ortony, 1977; Smith, 1978; Trabasso, 1980). Rumelhart and Ortony (1977), for example, in their investigation of schema theory, remarked that at a general level it was possible to regard "the entire comprehension process in schema theory as itself being a case of analogical reasoning" (p. 120). When the reader/listener determines that a particular situation fits a particular schema he/she has made a determination that the current situation is analogous to the ones from which that schema was derived. Also, when the reader/listener makes inferences about unobserved phenomena within a particular situation he/she is "assuming their existence by analogy" (p. 120).

Clearly, meaning does not reside in print (Smith, 1978; Spiro, 1980; Spiro & Tirre, 1980). What linguistic and non-linguistic information provides is only "a skeleton, a blueprint for the creation of meaning" (Spiro, 1980, p. 245). Meaning, in actuality, arises from the interaction between information and the recipient of that information - from the analogous relationship between what Smith (1978) terms visual and non-visual information. Therefore, the existing gap between what is heard/read and what is known must be closed by a measure of what can be broadly described as analogous reasoning.

Further, as recent research by Anderson (1977), Bransford (1979), Spiro (1980) and others have demonstrated, everyone does not process information in an identical fashion. Individual differences

in information processing do exist. Therefore, if as was stated, comprehension is a reasoning process, then the reasoning strategies readers/listeners employ in the processing of information may function as primary sources of individual differences in comprehension. As Davidson put it, through analogy the listener/reader seeks to "render abstract or complex information comprehensible" (Davidson, 1976, p. 160).

The consideration of analogous reasoning is not a new tool in the investigation of individual differences, however. Spearman (1904, 1927) as previously noted, felt that general intellectual ability, as he designated it, could best be measured through abstract reasoning tasks. The Raven's Standard Progressive Matrices used in the current study was constructed on Spearman's theoretical position. Recent studies have also examined analogical reasoning items in reference to individual differences in cognitive processing. Several of these current investigations - Goldman & Bisanz (1980), Snow (1979, 1980a), Sternberg (1977, 1980), Whitely (1980) - will be examined and their implications relative to predictive accuracy of linguistic and non-linguistic information described.

Perhaps some of the most influential research in abstract reasoning has come from Sternberg (1977, 1980). One of the principal goals of his research has been the development of process theories of various reasoning tasks. In his 1977 publication, Sternberg identified what he felt to be the component processes involved in analogous reasoning tasks. These component processes visually and verbally detailed in Chapter III include:

Encoding - the process of inspecting a term and discerning its attributes.

Inferring - the process of forming between-term relations.

Mapping - the process of developing relation across-terms.

Applying - the process of predicting the correct term from information given.

While Sternberg's research has laid a foundation for componential analysis of abstract reasoning tasks, his attention has been directed toward standard analogies of the A:B::C:D format. It would be informative to see how the various performance components identified by Sternberg would operate in non-linguistic analogies of a more extended format. Such extended analogies are employed in the Raven's SPM. Further, it would increase understanding of reading/listening as reasoning tasks if the component processes applicable in non-linguistic analogies could be applied to linguistic materials as well.

Snow (1979) in his analysis of psychometric tests through experimental methodology has concluded that any cognitive learning or performance task requires the individual to execute one or more performance programs. A performance program is regarded as "an organized assembly of information processing activities designed to meet, as efficiently and effectively as possible, the performance demands of the test or task at hand" (p. 5). Snow's notion of performance programs coincides well with the performance components suggested by Sternberg's research.

As Snow (1979) theorizes, task instructions and sample items play an important role in task performance because they assist individuals in assembling the performance program appropriate to the execution of the task presented. If a task is a familiar one,

then the appropriate performance program may already be assembled and the individual must only mentally call it into action. However, if the task is novel, then an appropriate assembly may need to be created. This creation may take place by bringing together parts of various performance programs available to the individual. This concept of performance programs parallels, to some extent, the interpretation of schema theory. In schema theory, when confronted with a particular task the individual would survey existing schemata in attempts to find one which suits, in terms of knowledge, the task at hand. If no schema is found to be completely suitable to the task the individual may scan schemata for useable information and may reorder a schema to incorporate the knowledge needed to complete the task. This consideration of schema-learning is expanded later in this chapter.

Basically, within Snow's (1979) and Sternberg's (1980) theoretical framework task performance is dependent on the efficient and effective processing of various cognitive components. Consequently, variance in task performance may result from any of the component steps, and each of the components may contribute differentially to successful task performance. The adequacy of the component steps, in other words, will determine the individual's success on any specific task item. If analogous reasoning problems incorporate the component processes of encoding, inferring, mapping prior to application (prediction) then an inaccurate prediction, according to Snow's and Sternberg's postulation, might result from difficulties at any or all the component steps.

In his explanation of mental-tests and their interrelations, Snow (1980a) also noted that the relatively unspeeded power tests may

necessitate more flexible control and conscious endurance on the part of the individual than tests of differing construction. The Raven's Standard Progressive Matrices, for example, which Snow (1979) describes as a novel, power test "may involve an especially high degree of adaptive assembly" (p. 7). However, the individual's performance on one item should influence performance of consecutive items. The task format should become more familiar and less novel as the test proceeds.

In addition, on constructing a multi-dimensional scale of between-test correlations Snow (1980a) observed that the tasks that fell more in the central region of the scale involved abilities that he regarded as more general, central or complex than those tasks that fell in the peripheral regions. Were the chart, as Snow speculated, a dartboard the bulls-eye would be occupied by the Raven's SPM and very similar tests. Snow concludes that the Raven's is an archetypical example of mental-tests that "represent to a greater degree the kinds of assembly and control processes needed to organize on a short term basis adaptive strategies for solving novel problems" (p. 12).

On this issue of adaptive strategies Snow (1980a) prefers to elaborate on the Cattell and Horn Crystallized-Fluid theory of intelligence. He takes this position because he considers existing cognitive models of information processing to provide too narrow a set of processing concepts and measures. This is especially true in the consideration of reasoning facility, particularly where adaptive strategies are required to perform novel tasks. The concept of C_f or fluid ability seems better able to serve as the flexible foundation for component analysis of reasoning abilities in Snow's view, more than current cognitive processing models. It speaks more to the difference between

performance in novel, learning situations and performance in situations where accumulated educational knowledge is the focus.

Very recently, several studies have appeared which demonstrate the application or variation of Sternberg's component model to various analogous reasoning tasks. Whitely (1980) for example has taken Sternberg's componential analysis model and developed from it multicomponent latent trait models - a quantitative method that permits the examination of the contributions of both item and person differences to component process outcomes. The outcome differences studied by Whitely (1980) and Whitely and Schneider (1980) included response time and response accuracy, and the process outcome model was applied to two types of aptitude items - verbal analogies and verbal classifications. The stated purpose of these investigations, according to Whitely, was to study aptitude test items as information processing tasks. Further, the perceived advantages of such an approach appeared to be two-fold. First, the investigation considered a task, aptitude test items, which is of general significance in cognitive functioning. Second, the study anchored aptitude tests to cognitive theory - a link between psychometric and cognitive psychology sanctioned by Snow, Sternberg, and others.

Whitely and Schneider (1980) employed the multicomponent latent trait model in analysis of the responses of 104 undergraduates to 35 verbal analogies and 45 verbal classification items. Verbal classifications were of the type:

Verbal Classifications	<u>Example</u>	dog elephant parrot hamster (stem)
	<u>Alternatives</u>	1) tree 2) flower 3) barn 4) cat 5) grass (ellipse)

Verbal analogies were of the standard A:B::C:D format and were presented under three treatment conditions.

Verbal Analogies	<u>Example</u>	Tree:Sap::Man: (stem)
	<u>Alternatives</u>	1) Axe 2) Woman 3) Maple 4) Blood 5) Arm (ellipse)
Verbal Analogies Subtasks	<u>Image Construction</u> - Subjects are asked to specify the rule or set of attributes for stem completion.	
	<u>Example</u>	Blue:_____::Yellow:Orange
	<u>Rule?</u>	_____
	<u>Event Recovery</u> - Subjects are given the stem and response-alternatives and asked to make any necessary changes in the rule required.	
	<u>Example</u>	Blue:_____::Yellow:Orange
	<u>Alternatives</u>	1) Purple 2) Red 3) Green 4) Brown 5) Gold
	<u>New Rule?</u>	_____
	<u>Response Evaluation</u> - Subjects are given the analogy rule and asked to select the response-alternative that fulfills that rule.	
	<u>Example</u>	Blue:_____::Yellow:Orange
	<u>Alternatives</u>	1) Purple 2) Red 3) Green 4) Brown 5) Gold
	<u>Rule</u>	(Color obtained when red is added to blue)

The verbal analogy subtasks were administered in two consecutive sessions after a delay of six weeks to minimize subjects' memory of

item content. The results of the study indicated that verbal analogies and verbal classifications differed from each other by the nature of the processing demands they made upon subjects. To illustrate, verbal analogies were found to require primarily "image" or rule construction for successful processing outcomes, and were therefore dependent on an inductive reasoning process. Verbal classifications relied more on both deductive and inductive reasoning according to Whitely and Schneider (1980).

Generally the results of the study provided support for the feasibility of modeling response accuracy from component process outcomes onto verbal aptitude items. Yet, though Whitely and Schneider did apply Sternberg's componential analysis model to verbal aptitude items their analogy items were still of the standard A:B::C:D structure and did not consider process outcomes over more natural, extended linguistic analogies. In addition, the study focused on verbal reasoning of undergraduates and, therefore, could not speak to the developmental aspects of reasoning ability.

In a study conducted by Cramer, Pascarella and Hendress (1980), however, there was an attempt to analyze certain relationships between reading comprehension and the analogical reading ability of children in grades four to eight. The investigators, building on the research of Sternberg and Whitely, first sought to construct a verbal analogies test incorporating easy to read words. The instrument, Verbal Analogies Test-Form 2 (VA-2), which employed nine analogy categories of six items each, was then administered to 307 children in 12 randomly selected classrooms in four schools. The students were all classified as low-

normal to high-normal ability students. This determination was made on the basis of the SRA Achievement Test; Level E-Form 1.

The findings of this study indicated a significant relationship between verbal analogy facility and reading comprehension. The data also showed a relationship between analogical reasoning and grade level. Yet, several cautions should be mentioned about these results. First of all, reading comprehension was determined by subjects' performance on several subtests of the SRA Achievement Test. Questions may be raised regarding the accuracy of such a test to discern reading comprehension in a natural rather than contrived fashion. Second, the investigation focused on analogical reasoning across nine discrete forms of verbal analogies such as functional, quantitative, and part-whole. The variability between these various analogy types, especially in light of their common A:B::C:D format, does not seem convincing. Finally, while the study did attempt to control for language ability by the use of easy-to-read words there was no similar control for differential background knowledge of subjects or consideration of possible effects of intelligence on the outcome of the data.

Goldman and Bisanz (1980) sought to investigate individual differences and developmental changes in children's ability to solve verbal and figural analogies - what is referred to in this investigation as linguistic and non-linguistic analogies. As Goldman and Bisanz observed, standardized aptitude and intelligence tests "capture reliable and systematic individual differences in cognition" (p. 1). But, these tests have demonstrated little impact on the learning process. Therefore, researchers, according to Goldman and

Bisanz should seek to identify the cognitive bases for observed performance differences. The successful completion of the analogies, whether figural or verbal, were viewed by the investigators as dependent on the subjects' ability to execute the processes of encoding, inferring, mapping and applying (predicting) and others accurately, efficiently and in the correct sequence.

From a survey of the literature on analogical reasoning studies and their own research Goldman and Bisanz made several valuable observations. First of all, there was evidence of developmental differences in analogical reasoning. For example, as children became older they were more likely to identify the relevant attributes of the analogy items spontaneously. Also they were more likely to establish more stringent criteria for judging the validity of an analogy. Sternberg and Rifkin (1979) in their investigation were also able to demonstrate developmental differences in analogical reasoning in regard to the strategies and speed of processing depending on the type of analogy employed.

A second point made by Goldman and Bisanz was that investigations into analogy problems should consider a variety of analyses and should attempt to discover differences in the information and criteria needed to solve these problems accurately. By having subjects identify the relevant stimulus attributes and/or establish the criteria for response decisions information about the task-relevant or task-irrelevant strategies employed becomes available. Data would seem to suggest that the proportion of subjects who use task-relevant strategies increases dramatically with age.

Finally, by way of general comparison between children who perform well and those who perform less well on analogy tasks, Goldman and Bisanz drew the following comparisons. Subjects in the eight-to-eleven year old group who performed well on the verbal and figural analogies: 1) inferred appropriate relations between the terms; 2) applied relational information to generate accurate completion terms; and 3) used available response-alternatives as information facilitating to their performance. By contrast those children who scored low on the analogies: 1) were less likely to infer appropriate relations between terms; 2) generated completion terms frequently inconsistent with relational information; and 3) were more susceptible to distraction from the response-alternatives provided.

All in all, the investigation of analogous reasoning problems - linguistic or non-linguistic, extended or standard - seems able to potentially furnish much knowledge about how individuals process information. Further, by the careful analysis of a variety of analogous reasoning items data pertinent to individual differences in cognition and comprehension perhaps can be obtained. It is especially important to investigate individuals' ability to reason abstractly if, as Thorndike (1917) claims, "reading is reasoning," or if, as Rumelhart and Ortony (1977) theorize, "the entire comprehension process in schema theory" is, itself "a case of analogical reasoning" (p. 120).

Summary

It would appear that reading, as well as listening, are in essence, reasoning processes. As such, they require individuals to infer

relations among otherwise separate, isolated bits of encoded linguistic information, and to map these "chunks" of given information onto knowledge schematically held. It would appear that only through the ability to apprehend relations among information, current and past, concrete and abstract, does that information become meaningful to the reader/listener.

Analogical reasoning tasks are one means of examining individuals' ability to discern relations among pieces of information. By analyzing individuals' performance on tasks of abstract reasoning greater understanding of the efficiency and effectiveness of the component processes required in such specific information processing tasks may be reached. Further, if the examination of analogical reasoning tasks can provide knowledge about the processing of information in general then it would seem that it should, in turn, provide knowledge pertinent to the processing of a particular form of information - language.

Reading and Listening

The comprehension of information would seem to be dependent upon many variables - the individual's cognitive abilities and learning style, the content of the information communicated and the purpose of that communication. Comprehension of information is similarly dependent upon the mode of communication - that is, whether information is transmitted by print or by speech. Though presentation mode plays a vital role in information processing, far too little is known about its potential effect on comprehension and cognition. It would seem that under certain conditions and for certain individuals, one manner of

information presentation might be advantageous over another. But it remains to be seen if "what" conditions and "which" individuals can be specified. Various studies have set out to explore the effects of reading and listening on the processing of information and several of those studies will be examined.

Sachs (1974), for one, employed reading and listening presentation conditions in her analysis of word and meaning retention. She was concerned with whether the pattern of retention was the same during reading as during listening. Consequently, the experiment was designed to consist of both an auditory and visual presentation component. In the auditory presentation 120 college students were read 24 passages of connected discourse. Of these 24, four functioned as warm-up passages and 20 as actual test passages. The passage varied in length from 27 to 180 syllables. Four differing amounts of intervening material were inserted between the passage and a series of test sentences. The test sentences were of the following four types:

Semantic - where the meaning of the original base sentence was altered by subject/object interchange, negation, or word substitution.

Formal - where meaning of the base sentence was retained though its form was altered.

Active/Passive - where the base sentence was changed from active to passive voice.

Lexical - where a synonym was given for a base sentence word.

Identical - where another reading of the base sentence was given.

A false recognition paradigm was employed with the subjects who were asked to indicate whether the test sentences they heard were "identical" or "changed" in relation to the original base sentence. Subjects were also told to rate the confidence of their decisions on a five-point scale. The procedure and materials for the visual presentation was identical to those used in the auditory presentation with the exception that subjects read rather than heard the passage and test sentences. Also, no subjects had received the previous auditory treatment.

Comparison of the data across presentation modes produced several interesting results. First of all, initial recognition was found to be higher in the listening mode though the relative amount of loss was higher for that mode as well. In addition, with the auditory presentation semantic changes were easier to recognize than others. While lexical changes were poorly identified. With the written material subjects recognized active/passage changes in test sentences almost as readily as they recognized semantic alterations.

Sachs (1974) also observed that within 80 seconds of interpolated discourse, subjects could not distinguish paraphrases from original passages sentences. Yet, retention of semantic content was high for that same period of time regardless of presentation mode. What Sachs found was a similarity in the general retention pattern of both auditorially and visually presented materials. However, Sachs postulated that in reading the only semantic cues that were available to the reader were the words and the linguistic form of those words. Whereas, in listening to discourse the stress and inflection may assist the listener in the comprehension of the message. Also by way of comparison Sachs noted that reading is a self-paced act while in

listening the individual processing speed is externally controlled by the speaker.

During the seventies several reading/listening studies were designed to assess good and poor comprehension of connected discourse in relation to presentation mode (Guthrie & Tyler, 1976; Kintsch & Kozminsky, 1977; Smiley, Oakley, Worthen, Campione & Brown, 1977; Sticht, 1972). The results of these investigations suggested certain patterns in regard to presentation mode. To illustrate, in those studies where certain subjects listened to a passage for a fixed amount of time and other subjects read the passages for the same interval of time (e.g., Smiley, et. al., 1977 and Sticht, 1972) two common findings occur. Good readers seem to comprehend the material equally well, whether that material was read or heard. Poor readers, on the other hand, comprehended less well in the reading than in the listening presentation mode. These results seem understandable in light of the difficulty with which poor readers encode written information - a difficulty that would perhaps be reflected in processing delays. Wilkinson (1980) refers to this type of fixed interval presentation procedure as a deadline design.

For those experiments in which subjects were permitted the necessary time to read and respond to information (e.g., Guthrie & Tyler, 1976; Kintsch & Kozminsky, 1977; Sachs, 1974) two patterns were identified. For adult subjects working under such a procedure referred to as a completion design (Wilkinson, 1980), little difference in comprehension was observed between reading and listening. On the other hand, those studies using children as subjects have shown that listening to completion tasks generally produce better memory

and better comprehension than does reading to completion tasks (cf. Guthrie & Tyler, 1976). Similar results are frequently sighted as rationale for the use of listening comprehension tasks as indicators of reading potential in diagnostic procedures. Additional research is needed to confirm or disconfirm these response patterns across reading and listening tasks.

Daneman and Carpenter (1980) in their attempt to devise a measure of working memory span that taxed both processing and storage functions also furnished some insights into the question of reading/listening effects on information-processing. For that first experiment in the study, four tests involving narrative passages and factual questions were given to 20 undergraduates, native speakers of English. Short narrative passages followed by factual questions were used in this study because the investigators felt that this type of questioning strategy "taps a skill of both educational and practical importance" (p. 452). The four tests included a reading span test, a reading comprehension test that asked factual questions and also assessed pronominal references and a traditional word span test. Results of the first experiment showed a significant correlation between word span and reading comprehension with the reading span task related to working memory capacity.

In their second experiment Daneman and Carpenter (1980) wished to examine working memory when information was spoken as well as written. The purpose was to determine the relation between reading memory span and listening memory span and their correlation with reading and listening comprehension. Measures used in the experiment included a reading and a listening span test and a comprehension test.

The results of the experiment conducted on 21 undergraduates - proficient, adult readers - demonstrated a high correlation between measures of working memory span and measures of reading comprehension. Also for this sample, listening span measures were found to be almost as good at predicting reading comprehension as were the reading span measures.

An important observation made by Daneman and Carpenter (1980) relative to the reading/listening topic was that the relatively high correlations between subjects' performance on reading and listening tasks seem to indicate "that a significant amount of individual difference is common to both reading and listening" (p. 460). This seems particularly true in regard to populations of adult, proficient readers similar to those employed in the investigation. For such individuals difficulties in reading or listening may arise from comprehension processes that tax the working memory rather than from encoding processes. This conclusion reinforces Sticht's (1972) earlier finding that visual encoding assessment within information processing theory may not identify a significant source of variance common to comprehension of spoken and written discourse. This would seem especially true in view of the fact that both the Daneman and Carpenter (1980) and the Sticht (1972) studies consider reading/listening processing differences among adult readers. The validity of these findings in relation to younger, less proficient readers/listeners requires further verification.

Some additional knowledge about the effects of reading/listening on the processing of information was obtained from an investigation recently conducted by Wilkinson (1980). The purpose of the study was to assess the interactive effects of communication mode, skill of the

individual and nature of the text. As Wilkinson surmised "even novice readers" may understand simple linguistic information "as effectively by reading as by listening" (p. 561). By comparison only proficient readers may understand information equally well under either a reading or listening mode. Wilkinson uses the word "understand" to incorporate the processes of recognition, comprehension and memory. In the study, efforts were made to assess the effects of these three component processes - recognition, comprehension and memory - on understanding under both reading and listening conditions.

Recognition was examined by subjects' perceptual identification of language symbols. Comprehension and memory were assessed by subjects' ability to respond correctly to questions based on passages just read or heard, where there was no access to the passages during the testing process. By employing this type of questioning strategy, the study, in Wilkinson's view, would assess a mixture of comprehension and memory. A correct response to passage questions would demonstrate that subjects were able to construct mental representations of the passage information, and had inferred a relation across sentences within that passage.

Two experiments were then conducted to examine these component processes in relation to presentation mode, skill of reader/listener and text information. In the first experiment Wilkinson focused on grade level and materials as factors in understanding. Subjects were 157 children in grades two through six. Materials for the study consisted of the following:

Who test - where subjects were asked to recall the names of three characters within the passage and to identify prominent referents.

When test - where events from passages were recalled and sequenced.

Where test - where subjects were asked to recall and place in the appropriate location three farm animals discussed in the passages.

Procedurally, the reading mode consisted of an oral reading of passage materials, while a look-and-listen method was used in the listening mode. The look-and-listen mode permits the subject visual access to the materials being heard. Both oral reading and look-listen are less common communication patterns in actual educational settings than silent reading and listen-only modes of information presentation. Further, the look-and-listen approach allows subjects dual sensory access to text rather than one - a fact that may color the implications of the experiment in relation to listening effects.

What the results of Experiment One did show, however, was that there were no sex differences for reading/listening within this particular experiment. In addition, understanding as measured by the various measures increased with grade level and, not surprisingly in view of the point about bi-modality just raised, the proportion of correct responses was greater for the look-and-listen than for the oral reading presentation.

From the data, Wilkinson postulated the existence of three listening/reading stages. In stage one children are generally slow accurate readers and there is, consequently, loss of comprehension and memory in the reading mode. In the second stage, however, occurring around grade four, reading rate for children closely approximates the normal speaking rate. By the third stage, according to Wilkinson,

children can read at a rate more rapid than the moderate rate at which an individual speaks. The rapidity of input would appear to have an effect on the understanding of information. This hypothesis was the focus of Wilkinson's second experiment.

In Experiment Two, subjects were 150 children of a single age shown to have normal or super reading ability. The When test was administered to subjects in an oral reading and simple listening mode. The results of the data indicated that as reading rate accelerated relative to the rate of listening children's comprehension and memory in reading declined relative to comprehension and memory in listening. Wilkinson, in sighting these results, raised several objections to these findings. First, it was felt that the correlation obtained could have resulted from uncontrolled variations in subjects' reading ability. Second, Wilkinson noted that this correlation would be expected for the events portion and not the sequence portion of the When test since the sequence portion results remained constant across reading speed.

Generally speaking, the Wilkinson study provided insights into several important questions regarding the effects of reading/listening on information processing. Not only did Wilkinson analyze the interactive nature of the cognitive components of recognition, comprehension and memory on understanding but she analyzed those components relative to factors of age, skill of subjects and forms of textual information. Also, Wilkinson considered the advantages and limitations of various presentation modes and procedures in the investigation of reading/listening effects.

Summary

It would appear that the comprehension of information relates not only to the skills and knowledge of the reader/listener and to the content of the information presented but to whether that information is orally or visually transmitted. Though the modality of presentation is seemingly important in the determination of its comprehensibility very little is known about the effects of listening/reading on the processing of information. The available research seems to suggest that for adult, proficient readers the mode of presentation has little differential effect on the comprehension of linguistic information. For younger, less proficient readers, however, listening-to-completion tasks seem to be slightly more preferable in the comprehension of discourse than reading-to-completion tasks. Also, a listening modality takes preference over a reading modality in deadline design tasks for those readers for whom reading rate does not equal or exceed the moderate pace at which an individual speaks. Considering the relatively small number of investigations focusing on reading/listening effects such findings require the support of additional research to be more definitive.

Comprehension and Prediction

Comprehension, or the interpretation of experiences, lies at the center of the reading process. This process of understanding is founded on the individual's ability to translate seemingly separate pieces of linguistic and non-linguistic information into some coherent pattern or design. Basically, comprehension results from interactions between the "top-down" and "bottom-up" processes of the cognitive system. The term

"bottom-up" is applied in comprehension theory to identify the knowledge that arises primarily from the textual input - knowledge that contributes to the understanding of that text. By comparison, top-down processes incorporated in an interactive model of comprehension refer to the experiences, skills, attentions and intentions of the reader/listener brought into play upon the textual input (cf. Rumelhart & Ortony, 1977). In the interactive theory of information processing, these top-down and bottom-up components operate in conjunction to produce meaning (e.g., Canney & Winograd, 1979; Carroll & Walton, 1979; Fredericksen, 1979; Smith, 1979).

Smith (1979) discusses these top-down, bottom-up processes, or "inside-out" and outside-in" systems as he calls them, in relation to visual and non-visual information. In the "inside-out" system of information processing, comprehension begins with highly discriminating, higher-level processes centered in the brain and ends with selective attention to particular portions of the visual information. In the "outside-in" system of information processing, however, comprehension begins with the visual information and ends with some representation or interpretation of that information within the brain. Further, Smith concludes that meaning arises from the interaction of visually- and non-visually derived information - that is, from the interplay of the "inside-out" and "outside-in" systems of information processing. According to Carroll and Walton (1979) the major virtue of a theory that emphasizes top-down or "inside-out" components is that such a theory focuses attention on the cognitive processes of reading "in which the ongoing apprehension of the total meaning of a text" acts to facilitate "the apprehension of particular parts of the text" (p. 234).

As Frederiksen (1979) observed, there is little disagreement in the literature with the assumption that comprehension of discourse involves both top-down and bottom-up systems. However, controversy still continues over the extent to which the comprehension process is essentially top-down or bottom-up in nature. Frederiksen suggests that more research into the interplay between top-down and bottom-up processes seems warranted in light of the existing controversy.

One key idea that has emerged from the recent research into comprehension theory, however, is this notion that visual information is not meaningful in and of itself. Meaning comes about only when the reader/listener infers some relation between textual input and schematically-held knowledge. It is through the process of inferencing that the reader/listener is able to link pieces of information together into some "causal chain of underlying conceptualizations" (Trabasso, 1980, p. 1). These inferences, according to Trabasso, are highly automatic and generally unconscious components in information processing that perform several functions. First, as Olson (1970) noted, intended meanings of words are frequently ambiguous and inferencing is needed to resolve lexical ambiguities. Second, inferences assist in bridging the relational gap between nominal and pronominal references - gap suggested by the earlier example:

The man killed the man.

The man killed himself.

Third, context for information processing is often inferentially established. To illustrate, "Mary had a little lamb" may take on variant meanings if the context inferred was a nursery rhyme, a meal,

or a birth. Finally, Trabasso demonstrates that inferences aid in developing a broader conceptual framework for the interpretation of information. The distinction between inferences at the contextual and conceptual levels of information processing referred to in functions three and four are the differences between what Frederiksen (1979) identified as text-based and schema-based inferences. Text-based inferences are those relations based primarily in the textual structures - relations between letters, words, phrases and so on. Schema-based inferences, by contrast, are those relations based principally in the prior knowledge structures of the cognitive system. A schema, according to Anderson (1977) is a theoretical knowledge structure containing informational slots subsumed under a more general idea. In schema theory knowledge is organized into a system of structures "embedded in more dominant and more abstract structures" (Canney & Winograd, 1979, p. 2).

Rumelhart (1980) has continued to modify and refine the concept of schemata which he describes as "the building blocks of cognition" (p. 33). Rumelhart notes that these schemata are employed in a variety of ways, including the interpretation of linguistic and non-linguistic sensory data. As prototypes of meaning schemata have been analogously compared to plays, computer procedures and theories. In relation to the comparison of schemata to theories, for example, Rumelhart states that the processes of comprehension can be viewed as being similar to hypothesis testing. Comprehension, it seems, hinges on the reader's ability to identify a configuration of hypotheses or, if you will, schemata that can account coherently for the information conveyed by

the text. Should the reader fail to locate such schemata or configuration the text remains basically disjointed or incomprehensible.

Within the context of top-down, bottom-up processing theories Rumelhart (1980) speaks of the "conceptually-driven" and "data-driven" systems of schema activation. According to Rumelhart, in "conceptually-driven" or expectation-driven processing a schema may be activated by some particular subschema which may, in turn, activate the appropriate subprocedures. By comparison, a schema is said to be "data-driven" whenever the activation of a subschema leads, in some fashion, to the activation of various higher level schemata.

If a reader/listener should fail to comprehend a particular piece of discourse Rumelhart speculates that this lack of understanding may be attributed to one or more of the following reasons relative to schema activation. First, the individual may simply lack the appropriate schemata needed to bring meaning to that discourse. A second explanation might be that the reader/listener did not instantiate the necessary clues from the discourse to activate the needed schemata cognitively available to him/her. Finally, the incomprehensibility of discourse may be attributable to the fact that the reader/listener activates schemata that may be inconsistent with the schemata intended by the transmitter of the discourse.

One highly unresolved though central dilemma in schema theory research, in Rumelhart's view, is the question of how learning occurs schematically. It is Rumelhart's opinion that there are basically three modes of learning possible with regard to schema theory. Elaborating on earlier research he labels these three modes accretion, tuning, and restructuring. Accretion, according to Rumelhart (1980) is perhaps

the most common form of learning and involves the gradual accumulation of knowledge relevant to pre-existing schemata. In essence, through accretion, no new schemata are established, existing schemata are only expanded.

Tuning, on the other hand, involves the modification or evaluation of existing schemata which cannot meet the demands of current experiences. This schema modification may take the form of alteration in the variable constraints and default values of a schema either by the upgrading, generalization or specialization of those variable constraints or default values. Finally, learning may take place through the creation of schemata - what Rumelhart identifies as restructuring. There are basically two ways new schemata may develop, patterned generalization and schema induction. Through patterned generalization learning occurs by analogy, as new schemata emerge by the copying of similar existing schemata. The idea of schema induction which relates to the repetition of certain "spatio-temporal configuration forms of schemata" is a much more troublesome notion (Rumelhart, 1980, p. 54). This operation requires the presence of some system sensitive to repetitions in configurations of schemata that do not parallel the schemata that exist at the time of an experience. Generally, while still itself an evolving concept, schema theory research seems to hold much promise for the greater understanding of comprehension and prediction.

Essentially, within the interactive model of information processing the foundation of comprehension appears to be the individual's knowledge base - knowledge centered in visual and non-visual information. For example, as Bransford (1979) stated, individuals must possess knowledge of language in order to elicit plausible expectations about information

which brings about clarification of meaning. Trabasso (1980) adds to this prerequisite list by noting that inferences require more than vocabulary knowledge of text structures. Trabasso posits that inferences also require background knowledge, knowledge of social interaction and human intentionality and knowledge of causal relations between events. The emphasis placed on these various requirements differs from situation to situation depending on the nature of the information to be processed and the purpose to which the information will be put. In all instances, however, comprehension necessitates that the individual is "inevitably an active agent in his [her] own learning" (Anderson, 1970, p. 349), and inferences are an inevitable component in that active processing (e.g., Bransford, 1979; Frederiksen, 1979; Hansen & Pearson, 1980; Smith, 1979; Trabasso, 1980).

If understanding is prefaced by inferencing then this understanding is also guided, examined and perhaps facilitated by prediction (e.g., Olson, 1970; Smith, 1979; Trabasso, 1980). Prediction or the elimination of unlikely alternatives (Smith, 1978) is a component of information processing that has not been widely explored. Yet, the process of prediction is one of the foundations of reading - the ability to ask the specific questions of the information that leads to the reduction of uncertainty.

These predictions, as Smith (1978) perceives them, are multiplex, and interlaced. Smith discusses two levels on the prediction continuum that have an effect on the processing of information - focal and global predictions. Focal predictions, at the lower level of the prediction continuum, are involved in more detailed, text-bound judgments. Smith postulates that these focal predictions occur more suddenly and are

of concern for shorter periods of time than their more global companions. On the other hand, global predictions involve more general intentions that bring readers/listeners toward their comprehension goal. As top-down and bottom-up processes are interactive so are global and focal predictions, which work in conjunction to bring about understanding of information.

Predictions, whether global or focal, should not be conceived as "reckless guessing" (Smith, 1978, p. 66). They are considered choices - decisions based on textual and schematic knowledge. Further, the reliance or emphasis placed by the reader/listener on textual and/or schematic knowledge may be as much a consequence of that individual's cognitive style of information processing as it is likely to be a consequence of the content of the information and the processing-objective (cf. Spiro, 1979). As Spiro (1979) observed from his research, there are apparent tendencies for readers/listeners to use more processes in a preferred top-down or bottom-up direction when the situation allows for flexibility in processing direction. In addition, Spiro postulated that for some children there may be a tendency to invoke a particular style of information processing regardless of the mode in which that information is transmitted and regardless of the type or content of that information. In regard to the relative importance of such a contention to information processing theory further research in this area seems warranted.

In their investigations into prose recall and readability, Miller and Kintsch (1980) have also contributed to the knowledge of inferences and predictions, or expectations as they term them. From their research, Miller and Kintsch have speculated that the global meaning of a text is

built upon a framework of macropropositions and macrostructures - larger, more meaningful informational and structural units. These macropropositions and macrostructures are developed from smaller units of discourse - micropropositions and microstructures - which are inferentially fused into broader, more global pieces of information. The processing of these structural and informational units, according to Miller and Kintsch, are facilitated by correct predictions or expectations on the part of the reader/listener. As evidence of the facilitating character of predictions, Miller and Kintsch point to reading speed data. Those segments of passages characterized by more accurate predictions of upcoming events were calculated to require shorter reading times than those passages containing a number of inaccurate event predictions.

Miller and Kintsch (1980) also noted that while a "sound expectation" should generally facilitate information processing, an "overly specific expectation will be a waste of time" (p. 14). This last assumption was based on the contention that the more specific the expectations made, the greater the likelihood of mismatch between expectations and text. In addition, Miller and Kintsch (1980) described three conditions under which they felt accurate expectations or predictions were more likely to occur. Those three conditions included: 1) abstraction from features; 2) linguistic signaling of text structure; and 3) intersection of real-world and text information.

Generally, what Miller and Kintsch's work in recall and readability has seemingly demonstrated is that text is made more readable by its organization into more global informational and structural "chunks" known as macropropositions and macrostructures - an organization made

possible through inferencing. Further, the reading process seems facilitated if readers are able to establish accurate expectations based on the knowledge already acquired from the information processed. This compliments Smith's (1978) observation that predictions facilitate information processing by relieving the visual system and memory from potential overload by reducing text uncertainty in advance of processing.

Fillmore (1981) has approached this idea of text comprehension and prediction from the integrated perspective of cognition and linguistics. His research has centered on the investigation of third and fifth graders' interaction with standardized reading comprehension test items. The objective of his study is to profile the "ideal reader" of such test items - items which Fillmore identified as "a new genre for English written language" (p. 4). Specifically, Fillmore lists three goals for his ideal reader - real reader research. The first of these goals is to analyze passages and questions to determine the background knowledge and the various interpreting and integrating skills they require of the reader. The second goal of the study is to determine (through probing and interviewing) whether particular subjects possess the knowledge and skills necessary to process such texts. Finally, as the third goal of this research Fillmore and his colleagues are seeking to understand the ramifications of this developing genre, the standardized reading comprehension test.

Fillmore cautions, however, that the notion of idealization he presents is more suitable to written language than to spoken language, since written texts are more monologic, closed - in a word, composed. Yet, should the spoken text also be composed, as in the current investigation, this notion of idealization could also be applied.

That is not to say that the text being employed in Fillmore's investigation is well-developed. As Fillmore observed, the text was found to be "seriously flawed" requiring "a degree of tolerance and cooperation" which the reader should not generally be expected to possess (1981, p. 4). Further, the methodology devised to compare some theoretical ideal reader to a particular real reader involves a rather lengthy procedure where the reader's expectations regarding the text are solicited segment by segment. This micro-analysis of the text may be adding an additional burden to the subjects. Consequently, Fillmore feels that the effectiveness of the methodology depends to a large extent on the warmth and charm of the interviewers and on the motivations the subject may derive from the one-on-one attention from an adult.

From his work Fillmore has come to the position that the principal dynamic aspect in the reading process is the reader's formulation and revision of what he terms Envisionments. An Envisionment, according to Fillmore, is some "coherent 'image' or new understanding of the states of affairs" compatible with the language of that text and arising from the set of possible understandings which may also be suitable. In addition, these Envisionments can occur at various confidence levels designated E^0 , E^1 , E^2 , and E^3 . Consider the following example:

The princess ate some jam. The queen slapped her.

The princess cried.

At the first or E^0 level of Envisionment the states of affairs are those which are explicitly justified in the linguistic material. For the given example the E^0 level would consist of three disjointed segments. The reader would discern only that (sentence 1) a princess consumed some

jam, (sentence 2) that some queen slapped some undesignated female, and (sentence 3) that a princess for some unstated and unconsidered reason cried. By contrast, at the E^1 or second confidence level the states of affairs arise from a cohesive or connected text where the princess in sentence one is perceived by the reader to be the same individual referenced by the "her" in sentence two and mentioned again in sentence three. Further, the queen presented in the text is assumed to be the mother of the princess. This combining or linking of sentences required at the E^1 level is the source of difference between it and the E^0 Envisionment level.

It is at the third level of Envisionment, the E^2 level, that the reader injects a sense of goal or motivation into the story or message. For narrative texts this situation might arise from the reader's knowledge of story structure or understandings of the goals and characterizations predominant in such story types. When the text is non-narrative the goals and purposes of the author might be applied in the interpretation. In relation to the example given, the reader may make sense of the story by assuming that the queen slapped the princess as punishment for the princess' having consumed the jam. The crying may then be attributed to the pain or guilt the princess felt as a result of having been punished. Finally, at the E^3 or highest level of Envisionment, the reader embellishes the actual text with "details not motivated by the text itself" (1981, p. 14).

What Fillmore saw as the practical consequence of these various Envisionment levels was the realization that individuals incapable of making the required inferences, the reading-between-the-lines, required by the E^1 and E^2 levels could not be a good, successful reader. The

reader who remains at the E^0 level of Envisionment never succeeds in processing text as a meaningful, cohesive whole. Similarly, readers operating at the E^3 level of confidence trust their own embellishments too much and add elaborations not intended by the text.

While Fillmore's concept of the Ideal Reader is still unfolding, it does provide insights into the knowledge and skills required of individuals interacting with the genre we call reading comprehension tests. Further, through the extensive interviews and probes this research should furnish information relevant to the various expectations and predictions readers make when processing this specialized yet important form of discourse.

Summary

Comprehension is an interactive process. Understanding may arise from the interplay of top-down and bottom-up systems of information processing - systems that correspond to visually- and non-visually derived information. It would appear that such an understanding may arise from various combinations of visual and non-visual information. For example, comprehension can result from the relation of visual-to-visual information, as in the construction of macropropositions and macrostructures, or from the relation of textual to schematic knowledge. Further, it seems that the processing of any information requires the interaction of top-down and bottom-up systems though the emphasis placed on these various systems may relate to the nature and purpose of the information, as well as, to the cognitive style of the reader/listener. Generally, it appears that understanding is prefaced by inferences and facilitated by the presence of accurate predictions.

Accurate predictions are signals to readers/listeners that their understanding of the information processed corresponds to the intent of that information. In essence, there has been a conceptual match between the expected and intended meaning of information.

CHAPTER III

METHODOLOGY

Purpose of the Study

This study investigated the nature of predictive accuracy across linguistic and non-linguistic comprehension measures to ascertain if:

1. a relationship between linguistic and non-linguistic predictive accuracy exists
2. the mode of presentation affects performance on linguistic tasks
3. a measure of non-linguistic comprehension relates to a linguistic measure of intelligence
4. differences in prediction performance are effected by developmental differences.

Subjects

Twenty third-grade and 20 fifth-grade students from a middle class suburban Maryland parochial school were the subjects in this study. All students enrolled in the third and fifth grade at that particular school served as the population from which ten subjects at each of the grade levels were randomly selected. A representative number of males and females were included in the sample, 18 males and 22 females, though no effort was made to equalize sex representation. Each subject received all treatments.

Materials

The materials employed in this study are described in detail in the order of their occurrence in the experimental procedure.

The Raven's Standard Progressive Matrices

To examine subjects' non-linguistic predictive accuracy the Raven's Standard Progressive Matrices was administered. The Standard Progressive Matrices or SPM is a test of abstract reasoning ability. The scale consists of 50 problems divided in five 12-problem sets (Sets A, B, C, D, E). The 12-problem sets are so arranged as to provide within- and across-set gradation . To illustrate, the first problem in each set is the easiest resolved and each consecutive problem becomes progressively more difficult. The SPM also is designed to provide across-set gradation by structuring each set of items around a progressively more difficult theme. For example, Set A in the SPM series represents a reasoning concept constructed to be more easily grasped than that which is represented in Set C or E (see Appendix D for examples of the Raven's problems). Problem A-1 in the scale functions as a practice item to assist subjects in understanding the nature of the task. Each scale item in the SPM is followed by a presentation of several response-alternatives, one of which is appropriate for item completion.

The SPM published in 1938 and then revised slightly in 1947 was constructed on the basis of Spearman's (1923) investigations into the nature of intelligence, and is designed to measure an individual's ability "to apprehend meaningless figures presented for his [her] observation, see the relations between them, conceive the nature of the

figure completing each system of relations presented, and, by doing so, develop a systematic method of reasoning" (manual, SPM, 1976). In addition, due to the nature of the information presented, the test is to be relatively little influenced by the culture or education of the individual, and the SPM is currently employed as a "culture-fair" test of intelligence.

In regard to these stated purposes, supporting reliability and validity data are supplied. In terms of internal consistency the majority of the nearly 30 studies given reported correlations of at least .90, with a mode of .91 (Banks & Sinha, 1951; Elley & Macarthur, 1962; King, 1963; Loranger & Misizk, 1959). Test-retest reliability data from the original studies (Foulds & Raven, 1948; Raven, 1948) produced correlations ranging from .83 to .93. However, a 1967 report on American high-school students (Tully, 1967) resulted in test-retest correlations ranging from .55 to .84. Additional test-retest data are contained in at least 20 papers summarized in manual information.

Under the area of concurrent validity of the SPM, studies involving English speaking children and adolescents reported correlations of the SPM to the Binet and Wechsler scales ranging from +.54 to +.86 (Banks & Sinha, 1951; Moran, 1972; Raven, 1948; review by Burke, 1958). Additional cross-cultural research seems required before conclusions as to the concurrent validity of the SPM in different cultural settings can be drawn. The content validity of the SPM as measured by internal consistency on the other hand, appears to vary markedly when different test items are analyzed (Hornowski, 1959; Kalena, 1967; Keir, 1949). Generally, the results of the reported studies support the overall structure of the SPM though item analysis does suggest a predominance of medium-difficulty

items and some need to shift item order (Banks & Sinha, 1951; Keir, 1949; Sheppard, Fiorentino, Collins & Merlis, 1968). Finally, the Raven's SPM claims to be a measure of general intellectual functioning, and, according to numerous factor analytical studies reported among the literature receives some support for that claim (e.g., Gabriel, 1954; Gittins, 1952; Macarthur, 1962, 1968; Nisbet, 1953) (cf. Manual for Raven's Progressive Matrices and Vocabulary Scales, Part Three, Section Seven, "Research").

Graduated Passages and Inferential Questions

To investigate subjects' linguistic predictive accuracy, two parallel inventories of graduated passages and accompanying inferential questions have been devised. These parallel inventories were designated as Inventory I and Inventory II, and each inventory consisted of 30 short test passages and 6 sample passages. Every subject received one inventory as a silent reading task and one inventory as a listening task. However, the order of presentation was counterbalanced so that half the subjects received Inventory I as the silent reading task and II as the listening task, while the other half received Inventory II as the reading and I as the listening task.

The 30 test passages were divided into six levels (Sets B, C, D, E, F, G) of 5 test items each organized for increasing set difficulty. Consequently, each inventory was designed to provide across-set gradation of difficulty, and each set contained a sample passage which was used as a practice and placement item. Further, every passage was followed by an inferential question and several response-alternatives from which the subject was to select. Samples of the passages and questions are exhibited in Appendix D. Both passages and questions were taken from

the Barnell-Loft Specific Skill Series Drawing Conclusions Levels B, C, D, E, F, and G where B corresponds to a grade 2 difficulty level, C corresponds to a grade 3, and so on (Boning, 1978). The appropriate difficulty of the passages for a specific grade level was determined in three ways:

1. Readability Formulas - Materials were subjected to two readability formulas - the Spache Readability Formula and the Dale-Chall Readability Formula.

2. Vocabulary Load - The vocabulary used in the passages was determined by a vocabulary list which was a compilation of the work of Carroll, Harris, Thorndike and others.

3. Field Testing - Pupils used in field testing were pre- and post-tested on the California Reading Test, Stanford Diagnostic Reading Test, Iowa Test of Basic Skills, and the Gates-MacGinitie Reading Test. Final decisions on published materials were made on the basis of pupils' performance on Specific Series materials and these standardized tests (Hallam, note 1).

Each passage was typed for the investigation on a 5 x 8 card and was followed by an inferential question and response-alternatives displayed on a separate card. These passage and question cards were presented to subjects in sets. Through this manner of material selection, it was felt that the ecological validity of the material was increased. Ecological validity is defined as a form of external validity concerned with the "generalization of the results to other settings or environmental conditions similar to the experimental setting or condition" (Huck, Cormier & Bounds, 1974, p. 258). As such, passages and questions for the reading and listening tasks

were taken, unmodified, from materials widely employed in the classroom settings. Consequently, the use of such ecologically valid materials allows the investigator in this study to get "reasonable extrapolations" from the research "to the classroom milieu" (Shulman, 1970, p. 376).

In relation to the Raven's SPM, as a measure of non-linguistic comprehension, and the reading/listening passages, as measures of linguistic comprehension several parallels in design should be noted:

Parallels in Material Design

<u>Number of Items</u>	<u>SPM/Passages</u>
	Sixty non-linguistic and sixty linguistic items (30 each in reading and listening tasks) were presented.
<u>Organization in Sets</u>	The <u>SPM</u> is organized in five sets of 12 items each. / The passage inventories were organized in six sets of five items each.
<u>Inclusion of Sample Items</u>	The <u>SPM</u> contains one sample item for the scale. / The passages contained one sample item per set.
<u>Gradation in Item Difficulty</u>	The <u>SPM</u> incorporates gradation within and across sets. / The passages incorporated gradation across sets.

Increase in Number
of Response-Alternatives

The number of response-alternatives in the SPM increases from six (sets A & B) to 8 (Sets C, D & E). / The number of response-alternatives in the passages increased from 3 (Sets B, C, D) to 4 (Sets E, F, G).

Procedural parallels will also be presented later in this chapter.

The Slosson Intelligence Test for Children and Adults

The Slosson Intelligence Test or SIT was administered to provide data on subjects' performance on a linguistic measure of intelligence. The SIT was selected for its brevity and for the fact that it is a linguistic intelligence measure capable of being given by teachers and administrators. The items contained in the SIT are similar to those presented in Stanford-Binet tasks, though the SIT has a higher ceiling level for adults and a lower base level for children than does the Stanford-Binet.

A test-retest reliability coefficient of .97 was obtained for the SIT on 139 individuals ranging in age from four to 50 years. Children under four were not retested and IQ's obtained for this group should be considered tentative in view of low reliability correlations. In regard to concurrent validity, a correlation of .92 was obtained for the Stanford-Binet L-M, and the SIT for 141 subjects ages four to 19. While no correlation was presented for the SIT and the Wechsler Intelligence Scale for Children, an average score difference of 7.7 for a sample of subjects between the ages of six and 16 was noted.

A fairly exhaustive review of the literature produced little data relative to correlations for the SIT and the Raven's SPM. The one correlational study found which compared the SPM to the Slosson was conducted by Alcorn and Nicholson (1972) employing 191 adolescents with IQ's assessed to be at 90 or below. Their study produced a correlation coefficient of $+.43$.

Construct Validity of Materials

As indicated, the focus of this investigation was on the predictive accuracy of subjects across linguistic and non-linguistic comprehension tasks. Consequently, it was important to present evidence that the theoretical variable which these tasks attempted to measure was, in fact, being measured. Therefore, in this study construct validity was examined through an a priori and post hoc analysis of subjects' performance on selected task items relative to those items' perceived degree of prediction-difficulty. The perceived degree of prediction-difficulty was determined by two independent raters, doctoral students in reading, who analyzed selected linguistic and non-linguistic comprehension task items according to criteria described in the following section. This criteria-analysis furnished a numerical indicator of the prediction-difficulty for the specified task items for each of the criteria. Then, linguistic and non-linguistic comprehension task items were designated as high or low in prediction-difficulty depending on their average percentile ranking across the four criteria.

Determining the degree of prediction-difficulty. The difficulty in prediction, that is, difficulty in the reduction of uncertainty,

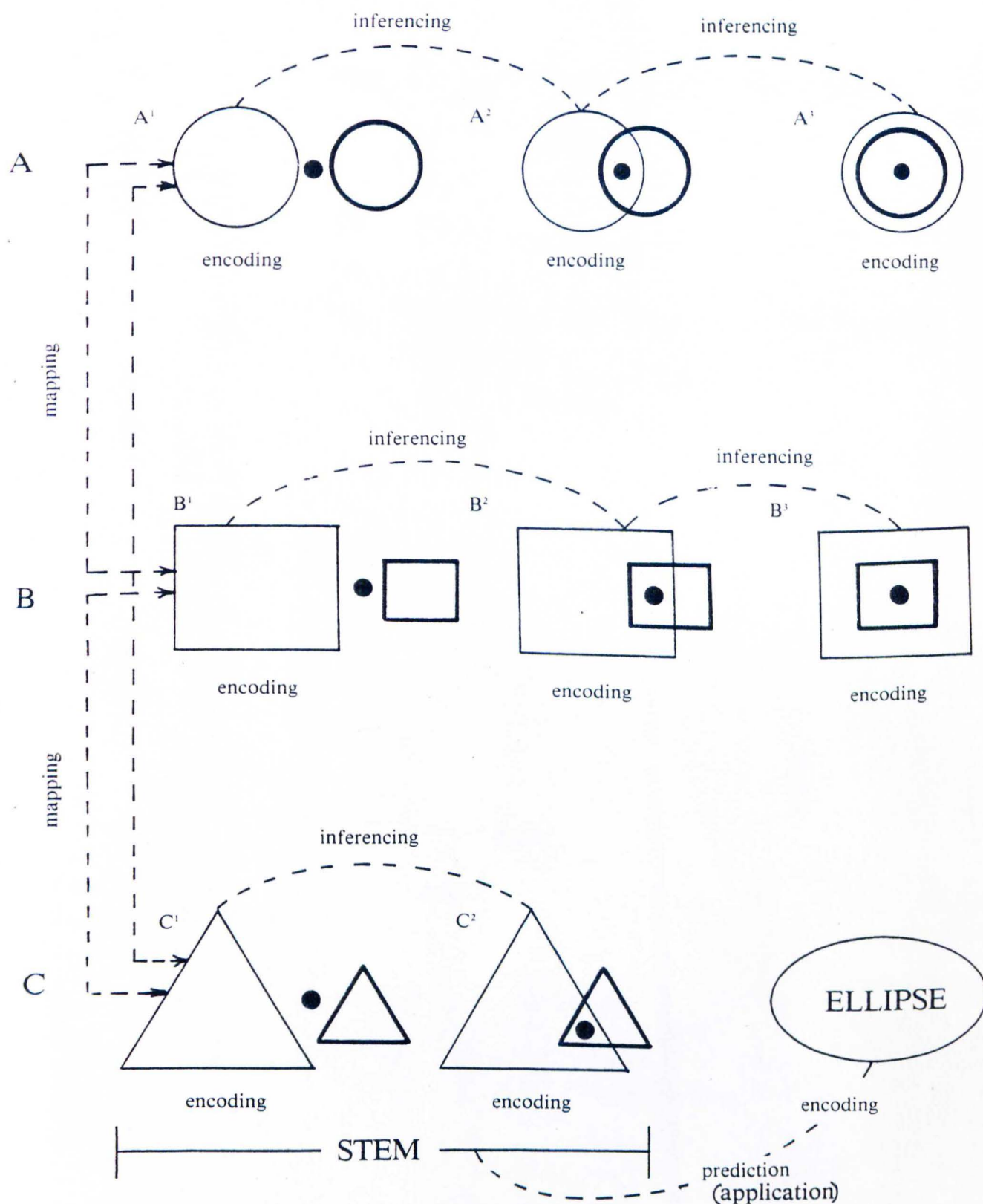


Figure 5
General Processing Model for the *Raven's SPM*
as an Extended Nonlinguistic Analogy

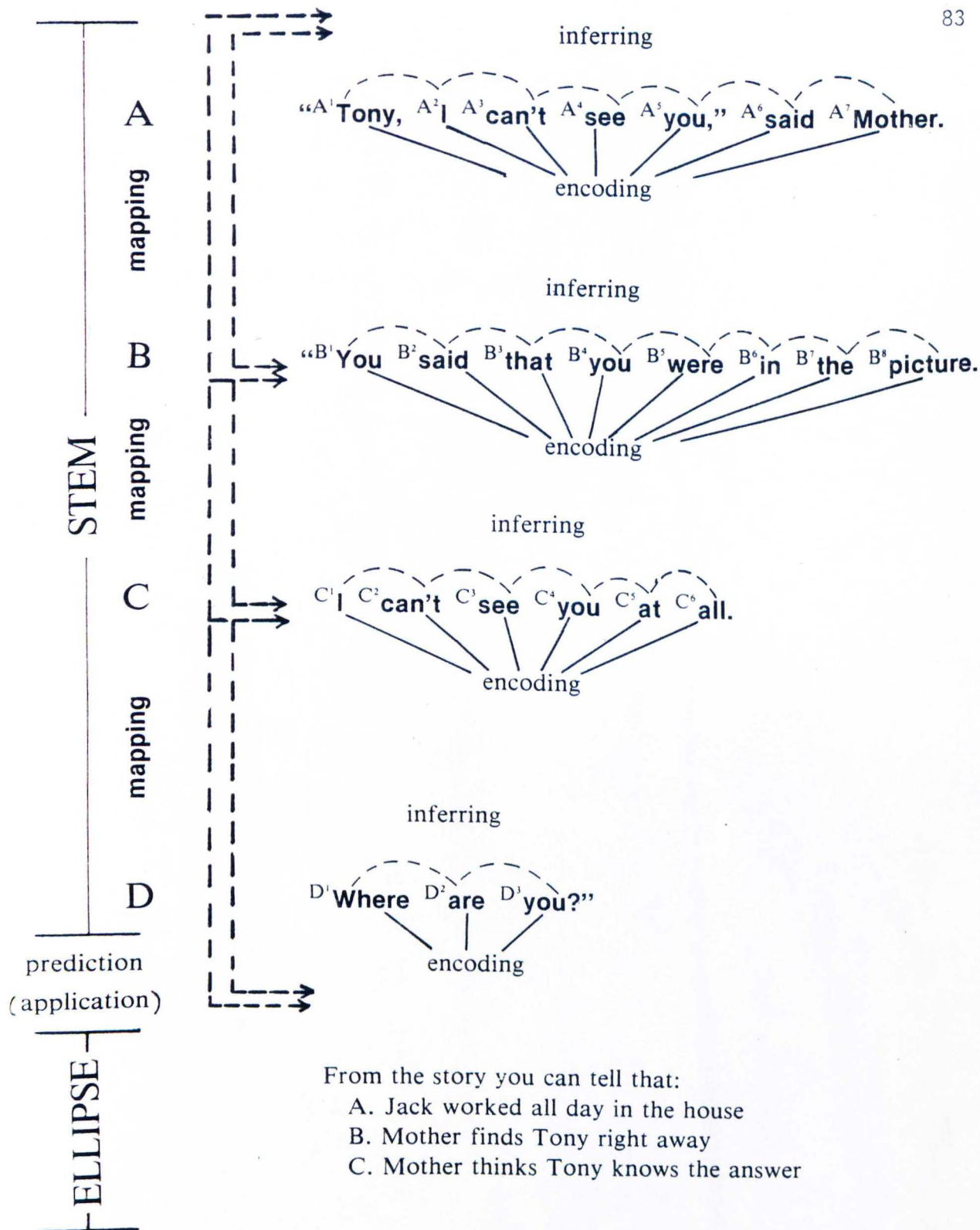


Figure 6
General Processing Model for Reading/Listening Passage
as an Extended Linguistic Analogy

for a particular subject for any piece of information will ultimately reside in a specific match between the subject's prior experiences and skills and the specific nature and content of the information presented. However, relative to the general processing involved in the linguistic and non-linguistic comprehension tasks under investigation in this study (see Figures 5 & 6), several criteria were described which would allow for the general evaluation of task-information in terms of the level of prediction-difficulty it would seem to require of subjects during processing. These criteria are paired in the tabular analysis according to the coding system in which they are structured - linguistic and non-linguistic. Further, the key processing-components of these criteria - terms, inferences, mappings, and encodings - have been defined and illustrated for clarification. These processing-component definitions arose from research in general processing theory (Spearman, 1923; Sternberg, 1977) and from recent research in analogical reasoning (cf. Goldman & Bisanz, 1980).

The basic unit to be processed, a term, may consist of one or more elements taken in conjunction. To illustrate, in the non-linguistic tasks, terms are denoted as figures composed of one or more lines held in conjunction. $\underline{A}^1$, $\underline{A}^2$, and $\underline{A}^3$ are three illustrations of terms (see Figure 5) presented in the non-linguistic comprehension task. For the linguistic task (see Figure 6) terms are denoted as words composed of one or more letters held in conjunction. Tony ($\underline{A}^1$), I ($\underline{A}^2$), and can't ($\underline{A}^3$) are illustrations of words as terms. It should also be noted that in these extended analogies, various terms are eventually held in conjunction to form metaterms. These metaterms are identified in Figures 5 and 6 by the large capital letters A, B, C, and so on, and permit the processing of larger chunks of information.

The internal representation of the terms of the analogy is the process of encoding where information is transmitted from the sensory system to thought.

Inferences are the relationships established between terms. It is through inferencing that terms $\underline{A}^1$, $\underline{A}^2$, and $\underline{A}^3$ in the non-linguistic example, and terms $\underline{A}^1$, $\underline{A}^2$ to $\underline{A}^7$ in the linguistic example are brought into conjunction, establishing the larger piece of information described as the metaterm $\underline{A}$.

Mapping is a process of forming relationships across terms of the analogy. To illustrate, in the extended non-linguistic analogy, a conceptual link can be made between metaterms $\underline{AB}$, $\underline{BC}$, and $\underline{AC}$. These conceptual links arise from mapping. Similarly, in the linguistic task, relationships between metaterms $\underline{A}$, $\underline{B}$, $\underline{C}$ and $\underline{D}$ can also be discerned through the mapping process. As demonstrated by Figure 5, a correct response to the question asked of the subject, requires the subject to map information from metaterm $\underline{A}$ onto $\underline{D}$ forming an $\underline{AD}$ link. Generally speaking, on the non-linguistic task, mapping will be required to process information across line while the linguistic task requires mapping across units that convey a complete thought. These units are frequently sentences though some compound, complex sentences may contain more than one thought-unit.

Consequently, the relative prediction-difficulty of the information presented in the linguistic and non-linguistic tasks in this study were determined by the evaluation of criteria presented in Figure 7. These criteria considered the relative difficulty of the processing-components just defined - processing-components from which the prediction process arises.

Figure 7

Criteria for Determining Degree of Prediction-Difficulty

<u>Linguistic</u>	<u>Non-Linguistic</u>
Number of words contained in the passage stem	<u>Number of Terms</u> Number of figures contained in the item stem
<u>Explanation:</u>	The more terms a stem contains, generally, the more difficult the prediction should be.
Average number of syllables contained in the words in the passage stem	<u>Complexity of Terms</u> Average number of parts to each figure contained in the item stem
<u>Explanation:</u>	The more complex the terms in a stem, generally, the more difficult the prediction should be.
Number of between-word relationships required in the stem	<u>Number of Inferences</u> Number of between-figure relationships required in the stem
<u>Explanation:</u>	The more inferences required in a stem, generally, the more difficult the prediction should be.
Number of across-thought unit relationships required in the stem	<u>Number of Mappings</u> Number of across-line relationships required in the stem
<u>Explanation:</u>	The more mappings a stem requires, generally, the more difficult the prediction should be.

A priori analysis. Two independent raters, doctoral students in reading, were asked to score ten items from the Raven's SPM and ten items each from the graduated passage inventories I and II. These raters were directed as follows:

1. Determine the numerical score of each task item for each of the specified criteria of prediction-difficulty. For example, calculate the number of terms presented in the item stem, and record it on the score sheet provided (see Table 1).
2. On the basis of the numerical value calculated, determine the rank-order of all items within a particular criterion separately for each of the tasks - Raven's SPM, Inventory I and Inventory II. For example, a selected Raven's item may receive a ranking of from 1 to 10 depending on its relative numerical position for criterion #1 (number of terms) where 1 is low and 10 is high.
3. After all selected items have been scored and ranked for each criterion, an average ranking for each item is determined. For example, if item #1 is ranked as 5, 6, 2, and 9 for each of the four criteria, the average ranking for that item would be 5.5.
4. When average rankings for all selected items within a particular task have been calculated, those five items with the highest average ranking will be designated as high in prediction-difficulty. Those five items receiving the lowest rank-order average will be designated as low in prediction-difficulty.

The results of the prediction-difficulty evaluation were applied in a post-hoc analysis of subjects' performance data. Those items calculated to be at a higher degree of prediction-difficulty should be calculated by subjects' performance to be at a lower degree of predictive accuracy. In other words, as prediction-difficulty increases, predictive accuracy should decrease. Such a result would provide evidence that the tasks as designed in this study are measuring the theoretical variable they purport to measure - predictive accuracy.

Pilot study - Results of the priori analysis. Two independent raters, doctoral students in Reading Education, were directed to rate ten items from the Raven's SPM and ten passages each from Comprehension Inventory I and II. Fifty percent of the items were selected from the first half of each measure and fifty percent were taken from the second half. The ratings were conducted according to the directions previously outlined relative to the criteria presented in Figure 7. For each of the four designated criteria, inter-rater reliability correlation coefficients on the calculated rank orderings were determined to be .9967, .68, .9989 and .74 respectively. Correlation coefficient for the average ranking for all four criteria was .933. Inter-rater reliability for the assignment of high (H)/low (L) prediction-difficulty designations to the selected task items was 1.00. The items scored and their respective high (H) or low (L) prediction-difficulty designations are displayed in Table 2.

Pilot study - Results of post-hoc analysis. Though a pilot-study reported later in this chapter had been conducted to examine experimental procedures and materials a second pilot was required. This pilot was

Table 2
Linguistic and Non-Linguistic Task Items
and Prediction-Difficulty Level

Measure	Item	Ranking	Measure	Item	Ranking
Raven's SPM	B5	L		E5-2	H
	B6	L		E6-4	H
	B7	L		E7-2	H
	B8	L		E8-1	H
	B9	L		E9-5	H
	D8	H	Inventory II	B12-4	L
	D9	H		B13-2	L
	D10	H		B14-3	L
	D11	H		B15-4	L
	D12	H		B16-2	L
Inventory I	B7-4	L		E10-1	H
	B8-3	L		E11-1	H
	B9-1	L		E12-5	H
	B10-2	L		E13-1	H
	B11-3	L		E14-4	H

conducted for two purposes. First, the pilot was to provide data for the post-hoc analysis of subjects' performance on selected task items - the analysis needed in the determination of the construct validity of the experimental measures. Secondly, the pilot was conducted to assess the experimental methodology as applied to a normally distributed population. The earlier pilot study obtained results from a depressed sample of poor readers markedly variant from the subjects employed in the present investigation. Subjects for the pilot were six third-grade and six fifth-grade students at a suburban Maryland parochial school. All subjects received all treatments. The scores, means, standard deviations and correlations from the pilot are displayed in Appendix A. According to the results of pilot study the experimental methodology was determined to be appropriate for a normally distributed population (see Tables A, B & C).

In relation to the post hoc analysis of subjects' performance on selected task items the findings supported the construct validity of the experimental measures. A t-test for related samples was used to assess the differences between subjects' performance on high and low prediction-difficulty items. The hypothesis under examination was that the mean for subjects' performance on the "low prediction-difficulty" items should be significantly greater than the mean for performance on "high prediction-difficulty" items for all comprehension measures. For the assessed items on the Raven's SPM a significant difference was obtained for subjects' performance on high prediction-difficulty and low prediction-difficulty items, $t_{10} = 3.16, p < .01$. Significant differences were also found for mean performance on the selected task items in both Comprehension Inventory I, $t_5 = 3.002$,

$p < .05$, and Comprehension Inventory II, $t_5 = 2.921$, $p < .05$. These data appear to support the construct validity of the experimental measures. A similar post hoc analysis was also conducted on data for all subjects in this study, and these data are reported in Chapter IV.

Procedure

The procedure followed in the present study is described in detail in the order of task occurrence. It should be stated, however, that prior to the initiation of the first task an unspecified portion of time was set aside for dialogue between the examiner and the subject. It was the purpose of the discussion to explain briefly to the subjects the nature of the study in which they were to participate. Further, it was hoped that this initial dialogue would contribute to a rapport between examiner and subject. Considering the extended nature of the procedures to be outlined, it was assumed that such a rapport was important in achieving maximum subject-performance.

All four tasks which constituted the experimental procedure for this study - the Raven's Standard Progressive Matrices, the graduated passages and questions under a silent reading condition, the graduated passages and questions under a listening condition, and the Slosson Intelligence Test - were administered by the investigator in a one-to-one setting. In addition, these four tasks were administered in two sessions of approximately 45 minutes each. Session I consisted of the Raven's SPM, and the graduated passages and questions presented under the silent reading condition. Session II consisted of the

graduated passages and questions presented under the listening condition and the Slosson Intelligence Test.

In relation to the Raven's Standard Progressive Matrices and the graduated passages and questions under silent reading and listening conditions several procedural parallels should be noted:

Parallels in Procedural Design

SPM/Passages

Administration

All measures were individually administered.

Administrator

All measures were administered by the investigator.

Directions

Directions for these measures were designed to employ parallel phrasing and emphasize parallel processes (Ex., Based on what
 see piece
 on what you hear choose the answer
 read
 question
 that best completes the pattern .

Time

All three measures employed a work-to-completion design (cf. Wilkinson, 1980). That is, all measures were administered untimed allowing students to work at their own pace throughout the tasks.

Practice Item

All three measures were initiated with a sample practice item which was used to explain the process involved in the task more fully.

Session I

The Raven's Standard Progressive Matrices. As specified on the SPM answer sheet information regarding the subjects' name, age, sex, and grade were recorded. Then, the test booklet was opened to the first scale item, A-1, and the subject was directed as follows:

Look at this. It is a pattern that is not yet complete. The pattern is followed by several pieces that might be used to complete the design. Based on what you see, choose the piece that best completes the pattern. (If subject responds correctly examiner proceeds with directions. However, if the subject does not provide the correct response the examiner may explain more fully the nature of the task through reference to the appropriateness or inappropriateness of the response-alternatives presented.)

Each of these pieces is the right shape to fit the space, but they do not all complete the pattern. All you have to do is point to the piece you think is the right one to complete the pattern. (Turning to item A-2) They are simple at the beginning and get harder as you go on. However, if you pay attention to the easy ones, the later ones will be less difficult.

There is no need to hurry. Take all the time you need but work carefully. You may begin.

As the subject began working through the SPM scale the examiner recorded the time on the answer sheet.

It should be noted that the directions for the task had been altered slightly from those contained in the SPM manual. These alterations retained the gist of the standardized instructions and were justified for two reasons. First of all, the standardized instructions were constructed for British speakers of the English language and seemed potentially linguistically awkward for young speakers of American Standard English. Secondly, the results of the SPM will be reported and analyzed in raw score form. There will be no attempt to convert these raw scores to percentile points as presented in the normed tables.

Once the subject had completed all 60 scale items of the SPM the examiner recorded the time of task completion on the answer sheet where indicated. At this point the examiner turned the test booklet back to the item C-9 and asked the subject to verbally respond to the following directive:

When you began this task you were told that these (indicating the items) were patterns that were not yet completed. You were asked to find the piece that best completed the pattern. When you are ready I would like you to describe, as completely as you can, the pattern that you found in this item (pointing to C-9). What did the piece have to look like to complete the pattern?

When the subject indicated to the examiner that he/she was prepared to respond, the examiner turned on a tape recorder and recorded the response. Subjects' initial response were unprompted. However, following the initial response the examiner may have sought clarification of some point made in that response (e.g., "When you said that things were getting closer did you mean ..."). The examiner may also have employed questions to solicit additional information not provided in the original response (e.g., "When you were figuring out the pattern, did you look down the page, across the page or both?").

Graduated passages and questions under silent reading condition.

To initiate the silent reading task the examiner entered the subject at the appropriate point in the reading passage inventory. It should be recalled that the subject received either Inventory I or II as the silent reading task as determined by the counterbalancing procedure. The point of entry was determined by subjects' grade level. To illustrate, all third-grade subjects were given the sample passage and question from Set B of the reading inventory - that is, at the difficulty point one grade below their present grade level. All fifth-grade subjects were given the sample passage and question from Set D - also at the difficulty point one grade below their current grade level. On being given the sample reading passage and question the subject was directed as follows:

Read this to yourself. It is a paragraph followed by a question and several possible answers. Based

on what you read, choose the answer that best completes the question.

The subjects responded orally by providing the letter of the answer they felt best completed the question asked. If that response was correct, the examiner gave the subject the five test-passage cards and five question-cards for that set. If the subject's response was incorrect the examiner used the sample passage and question to explain the process involved more fully. This was accomplished through reference to the appropriateness of the response-alternatives presented relative to the information contained in the passage. Third-grade subjects were then presented with Set B test passages and questions, and fifth-grade subjects were given test passages and questions from Set D of the inventory. Once the appropriate test-passage cards and question cards were given to the subjects along with the matching response card on which they recorded their answers (see Appendix C) they were directed as follows:

Write the letter of the correct response on the card in the same-numbered space. (Demonstrate). When you have finished one card, turn it over and go onto the next, but once you turn a card over you may not look back. There is no need to hurry. Take all the time you need but work carefully. You may begin.

As each subject recorded his/her answers on the response-card provided, the examiner simultaneously recorded that answer in the appropriate space on the Comprehension Prediction Inventory Sheet (see Appendix C). In this way, the examiner could immediately

determine the percentage of correct responses made by the subject at any level. If that percentage correct was below 80%, the examiner terminated the silent reading. If the subject achieved a percentage above 80%, the subject was given the next set of test passages and questions, and response-card and was allowed to continue with the task until such a time as the 80% criterion level was not achieved or all sets had been completed.

Under those circumstances where a particular set level was completed with the highest degree of accuracy and/or at a rapid pace the examiner may have selected to advance the subject to a non-consecutive set level. For example, the subject may have been moved directly from Set B to Set D. All levels in the inventory not attempted by the subject but below the level of termination were considered as completed with 100% accuracy. To illustrate, if the subject completed Set B with 100% accuracy, Set D with 80% accuracy and Set E with 60% accuracy, it was assumed that a score of 100% was achieved on Set C though Set C was not actually attempted.

One exception to the procedure outlined related to the necessity for post hoc data analysis described in the section on Construct Validity. All subjects were required to process the designated inventory items even though the level of termination had been reached below the point where those items occurred. However, only those designated items must have been completed and their results were incorporated only in the post hoc analysis for those subjects working beyond the point of termination.

Session II

Graduated passages and questions under a listening condition.

To initiate the listening task the examiner entered the subject at

the appropriate point in the listening passage inventory. This point of entry was determined by subjects' performance on the silent reading task. More specifically, the basic point was identified as that set level, one level below the point of termination in the silent reading inventory, unless that termination point was Set B. On presentation of the sample listening passage and question the subject was directed as follows:

Listen to this. It is a paragraph followed by a question and several possible answers. Based on what you hear, choose the answer that best completes the question.

Subjects orally provided the letter of the answer they felt best completed the question asked. If that response was correct, the examiner began the listening task. If the subjects' response was incorrect, the examiner used the sample passage and question to explain the process involved more fully. This was accomplished by analyzing the appropriateness or inappropriateness of the response-alternatives in relation to the information provided.

As subjects verbally responded to the passage questions heard, the examiner recorded that answer in the appropriate space on the Comprehension Prediction Inventory (see Appendix C). This permitted the examiner to determine the percentage of correct responses made by the subject at any particular set level. The subject was given visual access to the question cards. A criterion level and method was employed similar to that which has been detailed in the silent reading procedure.

Again, as described in the silent reading passage procedures, the examiner, under certain circumstances, may have made the decision to advance subjects to a non-consecutive set level. Further, those set levels not attempted by subjects but below the point of termination were analyzed as completed with 100% accuracy. Finally, those inventory items designated as part of the post hoc analysis must have been completed by all subjects though they were to be incorporated in the standard analysis procedures only for those subjects who had not surpassed the point of termination.

Slosson Intelligence Test for Children and Adults. The Slosson Intelligence Test was administered to all subjects in accordance with the standardized directions available in the test manual (see SIT Manual, 1977, p. 15).

Scoring Procedures

The Raven's SPM, the graduated passages and questions under a silent reading condition, and the graduated passages and questions under a listening condition were scored in similar fashion. Subjects' predictive accuracy scores for each of these measures was determined by the number of correct responses achieved on each. The selection of the correct response-alternative was scored as a +1 while the selection of an incorrect response-alternative was assigned no point value. For example, a predictive accuracy score of 25 on the SPM would indicate that the subject gave 25 correct responses out of a possible 60.

The scoring on the silent reading and listening tasks differed from the scoring of the SPM in that those sets unattempted by the subject on listening and reading tasks but under the point of termination

were counted as all correct. This added an additional five points to the subject's score for each set level skipped. To illustrate, a predictive accuracy score of 18 would have been achieved by a subject who answered five questions right in Set B, skipped Set C, answered four correct in Set D and four correct in Set E of Inventory I. This differed for the SPM where all items must have been attempted and then scored accordingly.

The scoring of the SPM and graduated passage questions under reading and listening conditions also differed in regard to the scoring of sample passage questions or items. Subjects' initial response to scale item A-1, which served as the procedural example for the SPM, was scored according to its correctness and incorrectness and the appropriate score value (+1 or 0) added to the score for the remaining 59 items. By contrast, subjects' performance on the sample passage questions under both the listening and silent reading conditions was not reflected in the predictive accuracy scores for those measures.

The Slosson Intelligence Test was scored according to standardized directions which yielded an Intelligence Quotient for each subject based on a comparison of mental age to chronological age (see SIT Manual, pp. 17-19).

Pilot Study - Procedure and Materials

A pilot study was conducted to determine the suitability of the experimental procedure and materials for this investigation. Subjects for the pilot study were six third-grade and six fifth-grade students enrolled in the University of Maryland Summer Reading Clinic during the summer of 1980. All subjects, undergoing remedial reading

instruction as part of the clinic program, received all treatments.

The total procedure was then monitored for the following:

- time spent on total activity and each task
- understanding of directions
- appropriateness of materials
- subject-examiner interactions

The investigator met with each of the subjects in one-to-one sessions over a three week period. The results of the pilot are presented in Appendix B, and furnished the following information relative to the points of concern (see Tables D, E, F).

Time spent - The total experimental procedure was observed to take approximately two hours per subject. The time required per task ranged from approximately 15 to 45 minutes. In general, the least amount was required by the Raven's SPM and the most for the silent reading task. The length of time required for the total procedure was found to be somewhat debilitating.

Directions - The directions established for the study were found to be suitable.

Materials - Though the materials were generally found to be appropriate for the proposed study it was decided that linguistic Inventories I and II should be reduced in length from 60 to 30 passages each. It was felt that this reduction in passage inventories would provide a partial solution to the time problem described. In addition, it was determined to counterbalance task presentation. In this way, differences that might arise on the basis of presentation

mode could not be attributed to the specific passages assigned to each task.

Interactions - Initial dialogue and questioning procedures were found to be suitable for the study.

The changes suggested by the pilot study were incorporated in this investigation.

Statistical Analysis of Data

Statistical analysis of data for this study are discussed by presenting the research questions posed. Each research question will be followed by a statement of the hypothesis to be tested and by the consideration of statistical procedures appropriate to the testing of that particular hypothesis.

Language/Thought

Research Question #1 - Is there a relationship between linguistic and non-linguistic predictive accuracy?

The investigator's speculation about the question was as follows:

Hypothesis #1 - There will be a significant, positive relationship between linguistic and non-linguistic predictive accuracy.

Statistical Procedure - To test the stated hypothesis 3rd and 5th grade subjects' performance data for the

graduated passages and questions under silent reading and listening conditions were analyzed in relation to performance data for the Raven's Standard Progressive Matrices. A multiple correlation technique was selected to provide tentative support or non-support for this research hypothesis. The selection of this statistical test was made on the basis of its usefulness in ascertaining relationships among variables.

Research Question #2 - Is there a relationship between a linguistic measure of intelligence and a non-linguistic measure of predictive accuracy?

The investigator's speculation about the question was as follows:

Hypothesis #2 - There will be a significant, positive relationship between a linguistic measure of intelligence and a non-linguistic measure of predictive accuracy.

Statistical Procedure - To test the stated hypothesis subjects' performance on the Slosson Intelligence Test and Raven's SPM were analyzed by means of the Pearson Product-Moment Correlation technique. The Pearson Product-Moment Correlation is a parametric technique that was selected to provide tentative support or non-support for the stated hypotheses.

Reading/Listening

Research Question #3 - Is there a stronger relationship between

listening linguistic predictive accuracy and non-linguistic predictive accuracy than between reading linguistic predictive accuracy and non-linguistic predictive accuracy?

The investigator's speculation about the question was as follows:

Hypothesis #3 - There will be a stronger relationship between listening linguistic predictive accuracy and non-linguistic predictive accuracy than between reading linguistic predictive accuracy and non-linguistic predictive accuracy.

Statistical Procedure - To test the stated hypothesis a partial correlation technique was employed to assess the relationship between subjects' performance on the non-linguistic Raven's SPM and the silent reading linguistic task when performance on the listening linguistic task was held constant. In addition, the relationship between subjects' performance on the non-linguistic Raven's SPM and the listening linguistic task was assessed with performance on the silent reading linguistic task held constant.

Comprehension/Prediction

Research Question #4 - Are there developmental differences in linguistic and non-linguistic predictive accuracy?

The investigator's speculation about this question was as follows:

Hypothesis #4 - There will be significant developmental differences in linguistic and non-linguistic predictive accuracy.

Statistical Procedures - To test the stated hypothesis 3rd and 5th grade subjects' performance data for the Raven's SPM and the two passage inventories presented under silent reading and listening conditions were analyzed by means of a t test for related samples. The t test was chosen for its appropriateness in considering the significance of two correlated samples.

Summary

Descriptions of subjects, materials and experimental procedures have been provided in detail within this chapter. Research questions, and the investigator's speculations regarding these questions have also been presented.

CHAPTER IV

RESULTS

The purpose of this study was to investigate the relationship of predictive accuracy to linguistic and non-linguistic comprehension measures. Several procedures, predominately multiple and partial correlation techniques, were used to analyze the data relevant to the specific research questions of concern in this investigation. Such correlational techniques were selected because of their usefulness in ascertaining relationships among the various outcome measures. Further, the nature of the data and the sample from which those data were obtained meet assumptions for the particular parametric techniques employed.

Quantitative Analysis of Data

The first research question of this study was as follows:

Is there a relationship between linguistic and non-linguistic predictive accuracy?

A multiple correlation technique was applied to assess the relationship between subjects' performance on the Raven's Standard Progressive Matrices, as the non-linguistic measure of comprehension, and subjects' performance on the passage/question inventories under silent reading and listening conditions, as linguistic measures of

comprehension. The means and standard deviations for all outcome measures appear in Table 3 (Data appear in Appendix F).

Table 3
Means and Standard Deviations of Outcome Measures

<u>SPM</u>		<u>Silent Reading</u>		<u>Listening</u>	
$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
33.275	10.146	12.225	8.613	13.15	8.694

The multiple correlation coefficient was calculated to be $R_{1.23} = .7227$, $F_{(2,38)} = 20.78$ $p < .001$. The multiple coefficient of determination, the proportion of total variance accounted for by the specified variables was determined to be $R^2 = .522329$. Therefore, this statistical analysis revealed a significant relationship between non-linguistic predictive accuracy as determined by the Raven's SPM, and linguistic predictive accuracy as achieved in the passage/question inventories. Further, over half of total variance of scores on the Raven's non-linguistic task, or 52%, can be accounted for by variance within the linguistic measures. The specific correlations of concern are reported in Table 4.

Table 4
Correlations for Outcomes
Measures

SPM	Silent Reading			
	.6959	Listening		
	SPM	.6157	Slosson	
		SPM	.3755	Listening
			Silent Reading	.6777

The second research question of this study was as follows:

Is there a relationship between a non-linguistic measure of predictive accuracy and a linguistic measure of intelligence?

A Pearson Product-Moment Correlation technique was used to assess the relationship between subjects' performance on the Raven's Standard Progressive Matrices and the Slosson Intelligence Test. As shown in Table 4 the correlation coefficient between these measures was calculated to be $r_{14} = .3755$, $F_{(1,38)} = 6.2375$, $p < .05$. When considered by grade level the correlation coefficient between the Raven's and the Slosson was assessed to be $r_{14} = .3413$, $F_{(1,19)} = 2.508$, $p > .05$, for third graders in this study and $r_{14} = .6601$, $F_{(1,19)} = 14.67069$, $p < .01$ for fifth-grade subjects. Consequently, statistical analysis for all subjects revealed a significant relationship between

the Raven's SPM, as a non-linguistic measure of predictive accuracy, and the Slosson Intelligence Test, as a linguistic measure of intelligence. However, when considered by grade level no significant relationship was found between the Raven's and Slosson for third-grade subjects, while a strong relationship between these measures was obtained from data of fifth-grade subjects.

The third research question of the study was as follows:

Is there a stronger relationship between listening linguistic predictive accuracy and non-linguistic predictive accuracy than between reading linguistic predictive accuracy and non-linguistic predictive accuracy?

A partial correlation technique was employed to assess the strength of the relationship of the non-linguistic measure of predictive accuracy, the Raven's SPM, with one measure of linguistic predictive accuracy, either silent reading or listening, when the second linguistic measure was held constant. The strength of a correlation is determined by its distance from zero to the point ± 1 . Therefore, the partial correlation reflecting the statistically larger non-zero relationship would be considered as portraying the stronger relationship.

For the first partial correlation subjects' performance on the Raven's was correlated to performance on the silent reading task with listening performance held constant. This partial correlation produced a coefficient of $R_{12.3} = .4809$, $F_{(1,38)} = 11.434$ $p < .01$. For the second partial correlation subjects' performance on the Raven's SPM was correlated to performance on the listening task while silent reading

performance was held constant. The second partial coefficient produced a coefficient of $R_{13.2} = .5166$, $F_{(1,38)} = 13.836$, $p < .01$. The multiple coefficients of determination for the first and second partial correlations were $R^2_{12.3} = .2312957$, and $R^2_{13.2} = .2669184$, respectively. The correlation between the Raven's and listening accounted for a slightly larger proportion of the total variance than did the correlation between the Raven's SPM and a silent reading.

Despite the partial correlation between the Raven's and silent reading performance being a lesser distance from zero than the partial correlation between the Raven's and listening performance, the difference between the two partial correlations was not large enough to be statistically significant, $t_{17} = .1472$, $p > .05$. The relative strength of these two partial correlations, therefore, requires further investigation.

The third research question of this study was as follows:

Are there developmental differences in linguistic and non-linguistic predictive accuracy?

To ascertain if there were developmental differences across linguistic and non-linguistic predictive accuracy measures, the performance of third-grade subjects was compared to the performance of fifth-grade subjects for each of the predictive accuracy tasks. The means and standard deviations for all outcome measures analyzed by grade level appear in Table 5. (Additional analyses of these data, by sex, can be found in Appendix E).

Table 5
Means and Standard Deviations of Outcome Measures
Reported by Grade Level

	<u>SPM</u>		<u>Silent Reading</u>		<u>Listening</u>	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
<u>Third Grade</u>	28.5	9.41	7.55	6.932	8.8	5.946
<u>Fifth Grade</u>	38.05	8.47	16.9	7.523	17.5	8.829

A t test was employed to analyze whether differences in performance for third- and fifth-grade subjects on the Raven's, the silent reading, and the listening tasks were of statistical significance. As demonstrated by the means for the linguistic and non-linguistic predictive accuracy measures, fifth-grade subjects performed consistently higher in regard to predictive accuracy than did the third-grade subjects. Statistical analysis revealed a significant difference between third- and fifth-grade subjects for each of these three measures. The smallest developmental difference was obtained for the Raven's Standard Progressive Matrices, $t_{38} = 3.374$, $p < .01$, while the largest difference was noted for performance on the silent reading passages, $t_{38} = 4.088$, $p < .001$. Differences between fifth- and third-grade subjects on the listening task were also calculated to be significant, $t_{38} = 3.655$, $p < .001$.

Post-Hoc Data Analysis for Construct Validity

As described in Chapter III subjects' performance on selected items from the non-linguistic Raven's Standard Progressive Matrices and linguistic Inventories I and II under silent reading and listening conditions were to be analyzed in relation to the construct they purported to measure - predictive accuracy. Subjects' performance on five items from each measure assessed in a priori analysis to be low in prediction-difficulty were compared to five items from each measure determined to be high in prediction-difficulty. Measures were to be considered valid in this regard if subjects performed those items identified as low in prediction-difficulty with significantly better accuracy than those items identified as high in prediction-difficulty.

Subjects' scores for each set of high and low prediction-difficulty items was calculated by means of a t test for correlated samples. Means and standard deviations of differences are displayed in Table 6.

Table 6
Means and Standard Deviations of Differences for
High and Low Prediction-Difficulty Items

<u>SPM</u>		<u>Silent Reading</u>		<u>Listening</u>	
$\bar{D}$	SD	$\bar{D}$	SD	$\bar{D}$	SD
1.375	1.111	2.325	1.456	2.4	1.428

Statistical analysis of these data reveal that for each of three measures of concern - the Raven's SPM, $t_{38} = 7.827$, $p < .001$; the silent reading task, $t_{38} = 10.101$, $p < .001$; and, the listening task, $t_{38} = 10.627$, $p < .001$ - there was a significant difference between high and low prediction-difficulty items. These significant differences, therefore, serve as confirmation of the construct validity of the differential predictive accuracy difficulty for the linguistic and non-linguistic tasks under investigation.

Qualitative Analysis of Data

Subjects' verbalization to the investigator's request to describe the pattern discerned in the Raven's Standard Progressive Matrices item C-9 at the conclusion of the measures administration were recorded. These responses were then coded as to whether they displayed a task-relevant or task-irrelevant strategy and as to whether the respondent used a one-directional, bi-directional or multi-directional approach to problem-solving. A task-relevant strategy was one which suggested that the subject, as directed by the task, attempted to identify some pattern in the item design. Task-irrelevant strategies, by contrast, were those that displayed no such attention to the task directive (i.e., "find the piece that best completes the pattern"). Task-irrelevant responses generally fell into three categories:

1. random choice ("I just guessed");
2. repetition of nearest term to ellipse ("I picked the one next to the empty space");

3. selection of a response-alternative which was not presented in the item stem ("This one, because I didn't see anything like it").

A "+" was given for those responses considered to be task-relevant while a "-" was given for those responses determined to be task-irrelevant.

In regard to directionality, a "U" was given to those verbalizations which demonstrated that the subjects either looked across the rows, or down the columns in determining their response. A "B" was given to those verbalizations which showed that the subjects looked both across the rows and down the columns, and a "M" was given those verbalizations which demonstrated that the subjects looked not only down the columns and across the rows but also diagonally or in some other appropriate fashion.

Twelve randomly chosen responses were scored independently by the investigator and a second rater, a doctoral student in Reading Education. Of the twelve responses, there was total agreement between the two raters on all but one of the responses for an inter-rater reliability of .957. Because of the high inter-rater reliability for these twelve responses the investigator felt confident about completing the remaining scores independent of a second rater.

The scoring of the subjects' verbalizations revealed the following information. First of all, of the 20 third-grade subjects, seven out of the 20 or 35 percent gave task-irrelevant responses to item C-9 compared to only two out of 20 or ten percent for the fifth-grade subjects. Second, of those subjects providing a sense of directionality in their responses - 31 out of 40 or 77.5 percent - 20 out of the 31

gave a verbalization that connoted uni-directionality for 64.5 percent. What was even more noticeable was that only two of those 20 who demonstrated a one-directional strategy spoke of looking down the column. The vast majority of the one-directional subjects spoke of looking across the line, as one does in reading. Further, only seven subjects, two in the third grade and five in the fifth grade used a bi-directional mode of processing as signified by their verbalizations.

Summary

The four hypotheses under test in the present study were examined in this chapter and statistically significant results were found relative to each of the stated hypotheses. Post-hoc analysis of subjects' performance on selected task items was incorporated to provide statistical support for the construct validity of the various measures. Further, qualitative analysis of subjects' mode of item resolution was also included to furnish additional data relevant to the research questions considered in this study.

CHAPTER V

SUMMARY, DISCUSSION AND IMPLICATIONS

Summary

As an investigation of predictive accuracy, this study examined the relationship of subjects' performance on a non-linguistic measure of comprehension to their performance on linguistic measures of comprehension presented under silent reading and listening conditions. All third and fifth graders at a suburban Maryland parochial school were the population from which 20 students from each grade level were randomly selected as the sample for this study. This sample was comprised of 18 males and 22 females. All subjects received all levels of treatment which consisted of four tasks administered in two sessions.

Session I included the administration of the Raven's Standard Progressive Matrices with the added verbalization of item C-9, and graduated passages and questions from the Barnell-Loft Drawing Conclusions Specific Skill Series presented in a silent reading mode. In Session II all subjects completed graduated passages and questions from the Barnell-Loft Drawing Conclusions Series presented in listening mode. The Slosson Intelligence Test for Children and Adults was also administered within the second session. The graduated passages and questions under the silent reading and listening conditions were contained in either Inventory I or Inventory II. These inventories were distributed in a counterbalanced format to ensure that half of the subjects would receive Inventory I as the listening task and Inventory II as the silent reading task, while half of the

subjects would receive Inventory I as the silent reading task and Inventory II as the listening task.

A multiple correlation technique was used to ascertain the relationship of the non-linguistic comprehension measure, the Raven's SPM, and the listening and silent reading linguistic comprehension measures. A correlation coefficient significantly different from zero was achieved in the comparison of subjects' performance data for these tasks. Through the use of a partial correlation technique comparison of the relationship between subjects' performance on the non-linguistic measure and either the silent reading or listening comprehension task was calculated while holding the effects of the second linguistic measure constant. Both partial correlations demonstrated that the relationship between the non-linguistic task and each linguistic measure were significantly different from zero. The difference in relative strength of these partial correlations, however, did not prove statistically significant.

The relationship between the Raven's Standard Progressive Matrices as a non-linguistic measure of predictive accuracy and the Slosson Intelligence Test, as a linguistic measure of intelligence, was examined by use of the Pearson Product-Moment Correlation technique. The correlation coefficient produced from this comparison was also found to be significantly different from the zero (or a non-relationship) position. Therefore, the results of this study furnished support for those hypotheses regarding the relationship between non-linguistic and linguistic predictive accuracy and between the Raven's Standard Progressive Matrices, the non-

linguistic measure of predictive accuracy, and the Slosson, a linguistic measure of intelligence.

With regard to the hypothesis on developmental differences a t test for related samples produced significant differences in the direction of the fifth-grade subjects for all three measures of predictive accuracy. These results supported the hypothesis that fifth-grade subjects would be more capable of accurate prediction on those measures included in this investigation than would third-grade subjects.

Limitations

There are several limitations to the present study that should be considered. The discussion and implications of the study should be viewed with the following limitations in mind:

1. Subjects were third- and fifth-grade students from a parochial school in suburban Maryland. Therefore, the results of this study may not be generalizable beyond individuals of the same grade, school system, and geographic location.
2. The nature of this research required the use of specific materials which were applied in ways or for reasons slightly different from that for which these materials were perhaps intended. The results of this study, in addition, should not be generalized beyond materials of similar design.

3. This investigation was conducted in a school setting but for experimental purposes the subjects were taken from their classroom to a different room. Whether or not such conditions interacted with the experimental procedure to produce the observed results can only be ascertained by further research.
4. All measures in this experiment were administered in a one-to-one session. Whether or not this situation may have contributed to the findings of this study cannot be determined without additional research.
5. The linguistic materials applied in this investigation were discrete passages followed by inferential questions. The cohesiveness and completeness of these passages and the appropriateness of the questions relative to the text may be of varied quality and suitability to the age and knowledge of the subjects involved.
6. For the listening task portion of this experiment, passages and questions were read aloud to subjects by the investigator. There was no control for expression and intonation in the oral reading of the linguistic material from subject to subject.

Discussion

Important factors relative to predictive accuracy surfaced in this investigation with regard to the areas of language and thought,

intelligence and cognition, analogous reasoning, reading and listening, and comprehension and prediction.

Language/Thought

At the outset of this investigation the view of language and thought as autonomous systems was established as the theoretical framework around which the research would proceed. As Chafe (1970) and later Olson (1970) and others have observed, meaning is at the core of communication. Yet, no message, in and of itself, conveys meaning. That meaning must be derived from an individual's understanding of the perceived or intended referent of that message - a process that requires the utilization of linguistic and non-linguistic knowledge. One argument Olson (1970) employed to substantiate his claim that meaning and an utterance are distinct was through the consideration of ostensive definitions. An ostensive definition, is described as that situation where an object is defined merely by pointing to or indicating the object in question. Certainly, under such circumstances the thought would appear to exceed the limited linguistic utterance. In the current experiment, the verbalizations of subjects revealed many instances of such ostensive definitions. When asked to describe the pattern of a particular item of the Raven's SPM, C-9, several subjects made responses of the type, "First, I saw this, then it changed to this, then to this.... So, I knew it had to be this one," pointing to various terms in the item stem and a selected response-alternative. The words "this", "this one" and "it" were words which were meant to convey the numerous, specific characteristics of each term or item pattern.

"Verbalization" is a word Chafe (1977) used to signify "those processes by which nonverbal knowledge is turned into language" (1977,

p. 41). This verbalization is a creative act and the assumption made that the final output need not be "uniquely determined by the initial nonverbal input from memory" (p. 41). The qualitative data furnished by subjects' verbalization of item C-9 suggested that even subjects who were able to complete the Raven's task with a high degree of accuracy may have lacked the verbal, language skills to describe fully and adequately the pattern discerned in the item stem. By comparison, there were those subjects who described in elaborate language some pattern or design that did not match the pattern expected by item generators and required for accurate item completion. This apparent pattern mismatch was reinforced by the provision of an inaccurate response-alternative.

This situation, if affirmed by further research, would qualitatively support the view of Macnamara (1977), Huttenlocher (1974, 1976) and others that one can think in the absence of language and one can language in the absence of thought. The experience, in this instance the processing and resolution of non-linguistic comprehension items, seems to exist separately from the act of verbalization. From the qualitative data available in this study no one-to-one correspondence between language and thought could be evidenced. Thus, this data would appear to furnish tentative support for the theoretical premise on which this study is based - the autonomy of language and thought.

One point of contrast that might be raised between the linguistic and non-linguistic measures employed in this study is with respect to the amount of non-visual information that can be employed in the resolution of the task items. Smith (1978) described this non-visual information as the "in the head" knowledge that the reader/listener

had regarding the message transmitted. Due to the novelty of the Raven's SPM it could be assumed that less non-visual information could be employed in working through that task than could be employed in the seemingly more familiar linguistic passages and questions. In part, to compensate for this imbalance and to provide more interesting information, the passages incorporated in the Barnell-Loft Drawing Conclusions Series described situations, events, or individuals which would be new and perhaps interesting to the majority of readers (see Appendix D). While this procedure does not eliminate the potential influence of prior knowledge in item resolution it might, perhaps, lessen that influence.

The Raven's SPM, by comparison, provides minimal opportunity for subjects unfamiliar with this measure to employ prior knowledge in item resolution at the outset of the task. However, as a consequence of the SPM's design there is the opportunity for subjects to apply knowledge gained from the resolution of earlier set items in the resolution of later, often more difficult, task items. Thus, as the subject moves through the Raven's SPM, the amount of novelty may decrease while the opportunity to employ non-visual information may increase.

In summary, qualitative data provided tentative support for the autonomy of language and thought, while quantitative data demonstrated the relationship of selected linguistic and non-linguistic comprehension measures with regard to the construct of predictive accuracy. As Huttenlocher (1974, 1976) stated, thinking may involve symbolic processes of some form but those symbolic processes may not necessarily be those of a linguistic code. In the present study, for example, the Raven's SPM utilized a non-linguistic code presented in geometric form. The graduated passages and questions, on the other hand, employed a natural language code.

Intelligence/Cognition

From the results of the study it was demonstrated that a moderate, positive relationship existed between the non-linguistic Raven's Standard Progressive Matrices and the linguistic Slosson Intelligence Test. This correlation coefficient between the Raven's and Slosson hypothesized and calculated to be significantly different from zero may appear surprising in light of various noticeable differences between these two instruments. To illustrate, the Raven's employs a non-linguistic code comprised of various geometric terms, while the Slosson uses a predominately linguistic code. Secondly, information in the Slosson is orally transmitted to subjects while the information contained in the Raven's consists of scaled items processed by the subject and designed for thematic repetitions. Such thematic repetitions are less apparent in the Slosson which consists of scaled items tapping diverse areas of academic studies, including mathematics, science, and history.

Perhaps the most marked distinction between these two measures, according to Snow's (1979) interpretation of them, is the distinction between crystallized and fluid measures of intelligence. The Raven's, by Snow's definition, is an archetypical example of a measure which assesses an individual's processing capacity in novel situations - a test of fluid intellectual ability. The Slosson, on the other hand, is more the example of an intelligence test which assesses acquired or accumulated knowledge. As such, it is more within the domain of crystallized intelligence measures.

Though apparently from such dissimilar foundations, a significant relationship was hypothesized and obtained for the Raven's and Slosson. This relationship was anticipated on the basis of similar significant relations that are produced between other seemingly diverse psychometric

measures of intellectual capabilities. Spearman (1923) theorized that this relatedness was attributable to the presence of the "g" or general factor of intelligence. However, a more recent explanation of such an interrelational phenomenon (e.g., Carroll, 1976; Snow, 1979; Sternberg, 1980) and the position assumed in this study was that psychometric instruments usually incorporate basic cognitive processes in their completion. As two tasks of inductive reasoning, the Raven's and Slosson, therefore, were hypothesized to correlate significantly on the basis of these shared cognitive processes. However, the delineation and examination of these overlapping cognitive processes are not within the realm of this investigation and require further research.

In this experiment the Raven's Standard Progressive Matrices considered by some to be a culture-fair measure of intelligence was analyzed as a comprehension measure through which subjects' predictive accuracy could be examined. This is a non-traditional application of the SPM and one which was designed to yield information relative to the componential cognitive processes required in task completion and related to the completion of specific linguistic tasks. This approach seems to coincide with Carroll's (1976), Cronbach's (1957) and Snow's (1980b) call to combine the concerns of cognitive theory with the knowledge of psychometrics. By such an in-depth consideration of the Raven's this study was able to consider the relationship of predictive accuracy of third- and fifth-grade subjects across linguistic and non-linguistic measures. In this manner, the present study linked, in a small way, the concern of cognition for understanding process with the products of psychometrics.

Finally, many of the attempts to delineate the various factors and components of the intellect have been examinations of its perceived structure upon which some theory of intelligence has been generated. By comparison, it was the stated purpose of this study, as with the current work of Carroll (1976) - factor analysis of intelligence rooted in cognitive theory - to begin the investigation of intelligence within a stated theoretical framework. This study attempted a component analysis of intelligence also rooted in cognitive theory. What information ensued, therefore, was interpreted and analyzed in relation to cognitive theory. This was not to dismiss the potential limitations of cognitive theory in explaining the nature of intelligence (cf. Snow, 1980b). Such limitations should be noted in the analysis of the data from this study and can be examined in future research.

In the current literature much attention within the area of intelligence has been directed toward sex differences (e.g., Holden, 1981). However, no significant differences in performance on the two measures which have been employed in diagnostic situations to assess intelligence, the Raven's and the Slosson, were observed in the present investigation. The means and standard deviations for the various outcome measures as analyzed by sex are presented in Appendix E.

Analogous Reasoning

The relationship hypothesized between the Raven's Standard Progressive Matrices, as a non-linguistic measure of predictive accuracy, and the passages and questions under silent reading and listening conditions, as linguistic measures of predictive accuracy, was found to be significantly different from zero. Such a relationship was

anticipated in this investigation for several reasons. First of all, reading, as described by Thorndike (1917), was theorized to be a reasoning task. Even more specifically, as noted by Davidson (1976), it is through analogous reasoning that readers/listeners seek to "render abstract or complex information comprehensible" (p. 160). As the Slosson and Raven's were speculated to correlate significantly, in part, because they were both viewed as inductive reasoning tasks incorporating core cognitive processes, so, too, were the Raven's SPM and the linguistic passages and questions considered to be related.

The core component processes that were seen to underlie these non-linguistic and linguistic measures were taken from Sternberg's (1977, 1980) research in componential theory. These componential processes were employed to demonstrate the parallel in information processing that may be incorporated in these tasks. The linguistic and non-linguistic predictive accuracy measures were conceived to entail the processes of encoding, inferring, mapping, and application (predicting) in task completion. Though Sternberg's componential theory has been applied to analogy problems, these analogies have been of the standard A:B::C:D format. In the present study, there was an exploratory attempt to apply Sternberg's component processes to the Raven's SPM items and to the linguistic passages in Inventories I and II as extended analogy problems.

It was felt that the application of Sternberg's research in this fashion would not only increase understanding of reading/listening items as reasoning tasks but would expand our knowledge of non-linguistic analogies, as well.

The results of this study would appear to support Snow's (1979) contention that any cognitive learning or performance task requires the

individual to execute one or more performance programs. Snow's notion of performance programs, which coincides with Sternberg's performance components, are organized information processing activities designed to meet the demands of a particular task. It would seem that such performance programs or components aid in making tasks comprehensible.

Snow (1979) further theorized that task instructions and sample items play an important role in the assembly of the appropriate performance program. Care was taken in the current study to provide task directives which might convey to some degree the parallel nature of the processes incorporated in the three predictive accuracy measures. Also, sample items were included in each of the three measures to furnish subjects with, perhaps, clearer understanding of the processes involved in task completion. These procedural parallels were provided to highlight for subjects the performance programs or componential processes that were hypothesized to underlie the linguistic and non-linguistic predictive accuracy measures.

The scrutiny of Sternberg's component processes in relation to test items, an important, special form of information, was undertaken by Whitely (1980) and Whitely and Schneider (1980). In the Whitely (1980) and the Whitely and Schneider (1980) study demonstration of the feasibility of modeling response accuracy from component process outcomes for the linguistic items was undertaken. This study explored the feasibility of such modeling within the realm of linguistic, as well as, non-linguistic items, as did the research of Goldman and Bisanz (1980).

Few of the studies of analogical reasoning displayed concern for developmental aspects of reasoning. The current study, however, was designed to furnish data relevant to this issue. For all predictive

accuracy measures administered in this investigation there were significant differences found in the direction of the fifth-grade subjects. These findings are similar to those of Cramer, Pascarella and Hendress (1980) who found not only a significant relationship between verbal analogy facility and reading comprehension, but found a significant relationship between analogical reasoning facility and grade level in favor of subjects in higher grade levels.

Goldman and Bisanz (1980) also noted developmental differences similar to those achieved in this investigation. As with the present study, Goldman and Bisanz (1980) found that older children were more likely to identify the relevant attributes of the analogy items, and were more likely to establish more stringent criteria for judging the validity of the analogy as evidenced by the task-relevant or task-irrelevant strategies they employed. The qualitative data from this study pertinent to task-relevant and task-irrelevant strategies, suggests support for Goldman and Bisanz's contention that the proportion of subjects using task-relevant strategies in the resolution of analogy problems increased noticeably with age.

Reading/Listening

The investigator's hypothesis regarding the strength of the relationship between the non-linguistic task, the Raven's Standard Progressive Matrices, and the linguistic silent reading and listening tasks was not statistically supported by the data. That is, the relationship between the non-linguistic measure and the listening linguistic task was not found to be statistically stronger than the relationship between the non-linguistic measure and the silent reading

linguistic task when the second linguistic variable was held constant. Certainly, as demonstrated by the hypothesis this situation was anticipated. Such a position was taken, however, because of the direction of findings of previous reading/listening research which demonstrated that with children working within a completion design (Wilkinson, 1980) presentation within a listening mode was slightly preferable to presentation within a reading mode. The data of this study do not provide unambiguous support for conclusions of Guthrie and Taylor (1976) which stated that listening to completion tasks for children generally produced better memory and better comprehension than reading to completion tasks.

In the current investigation as with the Wilkinson (1980) study subjects' understanding of the linguistic materials under reading and listening conditions was assessed by subjects' ability to respond accurately (i.e., to provide the response-alternative expected by the task generator) based on passage information read or heard when no access to the passages was given during testing. As Wilkinson (1980) stated, within a questioning strategy subjects' accurate predictions would seem to demonstrate their ability to construct mental representations of the passage information and to infer relations across sentences within that passage.

In addition, the Wilkinson (1980) study showed no sex differences with regard to the reading/listening tasks. The present study, similarly, found no significant differences with regard to sex for either the listening task or for the silent reading task. Means and standard deviations for the specific outcome measures are reported in Appendix E. The variable of sex would not seem to be one which could explain

variability in presentation mode for the materials and samples employed in this investigation.

One important aspect of this study which requires discussion is the notion of listening which has been employed. In this research, as is often true with studies designed to allow for comparison of a listening vs. a reading mode of presentation, the interpretation of what constitutes listening is a limited one, designating more accurately an oral reading mode. Outside the sphere of reading, the term listening might be more correctly applied to a conversational, narrative manner of information conveyance. Tannen (Note 2) recently elaborated on the distinction between what linguists might identify as text and talk in the analysis of discourse. Text, generally speaking, is a composed, more formal presentation of information while talk is used to designate a more informal, conversational presentation of information. According to Tannen's interpretation, consequently, the connected discourse contained in the inventory passages, whether transmitted in print or through the air, is more characteristic of text than talk. Talk, as the more conversational pattern of information transmission is, perhaps, more naturally conveyed in a spoken modality and more naturally received in a listening mode than text would be, though talk as well as text can be recorded in print. Future research in the area of reading/listening may seek to examine this issue of text vs. talk by comparing subjects' performance on printed, text materials with performance on spoken, talk materials.

Comprehension/Prediction

The primary concern of this study was to examine the accuracy of prediction of subjects within the realm of specified linguistic and

non-linguistic comprehension measures. This examination was conducted through an assessment of the relationship of these specific outcome measures by means of a multiple correlational technique. The resulting multiple correlation coefficient provided verification of a relationship between the Raven's SPM and the Barnell Loft Drawing Conclusions passages and questions that was significantly different from zero. Certainly, such a relationship was hypothesized, in part, because it was felt that these tasks would demonstrate the interplay between the "on the page" symbols, whether those symbols were linguistic or non-linguistic in nature, and the "in the head" reasoning skills. This interplay corresponds to the data-driven, conceptually-driven; top-down, bottom-up; or, inside-out, outside-in models of information processing that operate in an interactive fashion.

These information-processing models incorporate the notion of the inference as a link between what is known and what is seen or heard, and as a link between "on-the-line" and "between-the-line" information. As Trabasso (1980) has observed, comprehension is prefaced by inferences and is guided, examined and perhaps facilitated by prediction. Within the particular genre in which it was analyzed for this study, the comprehension test item, the focus on prediction had been converted to an emphasis on the accuracy of prediction. Fillmore (1981) described the importance of the processor of information within this genre to match the expectations (predictions) of the generator of that information. Consequently, when the special context for information processing is a testing genre, accuracy of prediction may be considered preferable to prediction of a less specified outcome.

In the present investigation, the prediction under assessment could be described by Smith's (1978) terminology to be more global than focal in nature. Therefore, this study in considering the accuracy in predictions of the test processor to those predictions anticipated by the test generator were, in essence, researching the conceptual match between the sender and the receiver of specified information. Further, these predictions are multidirectional in form. That is, the questions asked elicit responses relative to what the readers/listeners assumed had transpired, as well as what they assumed would transpire. From the data it can be stated that those who were likely to provide the expected response-alternative for the Raven's Standard Progressive Matrices items were also more likely to furnish the expected response-alternative for the linguistic passage questions under both listening and silent reading conditions. In summary, Fillmore's (1981) notion of an ideal reader as one whose knowledge and expectations matched that required of a particular test item has relevance to present study. Further research in the conceptual match between test generator and processor beyond the relational consideration given it in this investigation seems warranted.

Implications

Theory

In this study, the qualitative and quantitative data have, it would appear, tentative implications for theory in three domains - language and thought, the structure of the intellect and cognition. With regard to the issue of language and thought autonomy subjects' verbalizations through a selected item of the Raven's Standard Progressive Matrices

appeared to substantiate the theoretical claim for the distinct nature of these systems. No one-to-one correspondence was observed between subjects' facility in item resolution and their facility in verbalizing the resolution process. Also, many subjects chose to verbally describe rather complex terms by means of ostensive definitions, thus allowing the object or term to describe itself. For those subjects and for those materials included in this research "knowing how" did not necessarily mean "telling how". Such results, however, require further investigation, and these investigations may wish to examine more intensely the parallel between performance ability and verbalization facility.

Another theoretical domain considered within this research was the concept of intelligence and the structure of the intellect. The findings of this study, as previously noted, suggested the potential feasibility of analyzing psychometric task items by means of component, or perhaps factor theories of intelligence. In this manner, the psychometric emphasis on product may be tempered or combined with a concern for the process from which an expected response may arise. The ability of psychometric instruments to predict future academic performance should be augmented by the examination of cognitive processes presently at work in bringing about successful task completion. Future research should continue this investigation into the processing demands and informational content of psychometric measures, especially those that play a greater role in the placing, labeling or programming of individuals.

Prediction, as a guide and facilitator of understanding, would appear to be an important though little understood process in comprehension theory. In this study a relationship between linguistic and non-

linguistic measures of comprehension was analyzed by means of subjects' accurate predictions. Such an accurate prediction was apparently predicated by subjects' ability to encode the necessary information, draw the expected inferences, and create the required mappings. The view of prediction seems somewhat analogous to the detective's procedure where the scene is surveyed (encoding), clues are collected (inferences) and conclusions as to what occurred are drawn (predictions). Such a basic reasoning strategy from the work of Rumelhart (1979), Spiro (1980) and others, would seem to be a necessary component of understanding. Parallels must be drawn between what is schematically-known and what is transmitted in the way of information and conclusions made. As research into comprehension continues, especially within a cognitive framework, careful examination of the role of prediction will need to be undertaken.

Research

The results of the current study have indicated the need for further research to support the findings presented within the areas of language/thought, intelligence/cognition, analogous reasoning, reading/listening and comprehension/prediction. The following are directions and rationale for further studies:

1. Limitations placed on the generalizability of the results call for investigations extended to include different age/grade-level subjects in both public and private schools.
2. In order to make the results of the study more generalizable, additional research will need to

apply a similar experimental procedure to varied psychometric and instructional materials.

3. While the present study sought to establish a relationship between specified linguistic and non-linguistic measures future research may wish to examine more complex questions with regard to this relationship. Appropriate statistical procedures will need to be identified which can address potentially more complex issues of predictive accuracy.
4. As noted, the definition of listening applied in this study was rather a limited interpretation applied to the reception of orally-read text. To provide a fuller picture of the differences between reading/listening future studies may wish to examine performance differences which might arise from the processing of information in text/talk modes of presentation.
5. The selected passages and questions employed in this investigation were chosen principally because of their wide application in classroom settings and not for the cohesive and complete nature of their discourse. Later research may wish to use linguistic materials of more adequate, literary quality that employ more appropriate, inferential questions.

Instructional Practice and Diagnosis

The possible implications for educators as indicated by this study are consistent with many instructional and diagnostic practices currently in use or can be regarded as supplementary practices to those commonly employed. These classroom practices are grouped for the purpose of discussion under the headings of instructional practice and diagnosis.

Instructional practice.

1. If as Thorndike (1917) hypothesized and as the results of this study seem to demonstrate reading is a reasoning task, classroom instruction should include the explanation of reading as a problem-solving activity. Students should be introduced to the fact that meaning does not exist "on the page", in many instances, but arises frequently from an interaction between knowledge contained "on the page", "between the lines" and "in the head".
2. This investigation demonstrated the application of analogous reasoning tasks as tools for analyzing reasoning strategies. Students should be introduced to analogies both of linguistic and non-linguistic varieties with discussion as to their means of resolution. Analogies may also be formulated by students and their rule for formulation discussed.

3. In the reading/listening setting instructors may preface new or novel information with an analogous situation or event with which students are perhaps more familiar. This type of technique has recently been tested by Hayes and Tierney (1980) with positive results obtained for the enhancement of comprehension through such an analogy-prefacing approach.
5. As the data from this investigation would seem to indicate, prediction is an important element in the comprehension process. Instructional practices should, therefore, introduce the idea of prediction to students, its meaning, and its value throughout the entirety of each reading experience. Opportunities to practice predictions with some follow-up discussion should be provided in a variety of settings such as a pre-reading, during-reading or after-reading activity.

Diagnosis.

1. When assessing students' responses to task items, it would appear from the findings of this study that more information regarding the response process can be garnered by the examination of subjects' approach to item resolution. Having subjects' verbalize their method of item resolution was the technique employed in this study.

2. Assuming that subjects, as evidenced in this research and the research of others (e.g., Rumelhart, 1980; Rumelhart & Ortony, 1977), must be able to employ reasoning strategies in the comprehension of oral or written discourse, it may prove informative in a diagnostic situation to assess such reasoning skills through the application of analogical reasoning tasks such as the Raven's Standard Progressive Matrices.
3. In the diagnostic setting it may serve the examiner to employ a measure of reasoning assessment less contaminated by accumulated knowledge or crystallized abilities. A non-linguistic measure of comprehension such as the Raven's permits the diagnostician to assess more fluid abilities, to apply Cattell's and Horn's terminology, with limited interference from language or culture.

Summary

In this chapter a summary of the present investigation has been provided with a discussion of the relevant findings within the areas of language/thought, intelligence/cognition, analogous reasoning, reading/listening and comprehension/prediction. Implications of these findings for theory and research have also been outlined, along with suggestions for instructional and diagnostic practice.

Reference Notes

1. Hallam, C. Personal communication, January 19, 1981.
2. Tannen, D. Analyzing discourse: Text and talk. Opening remarks at 32nd annual Georgetown University Round Table on Language and Linguistics, March, 1981.

References

- Alcorn, C. L., & Nicholson, C. L. Validity of the Slosson Drawing Coordination Test with adolescents of below-average ability. Perceptual and Motor Skills, 1972, 34, 261-262.
- Allington, R. L. Are good readers and poor readers taught differently? Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.
- Allington, R. L. Diagnosis of reading disability: Word prediction abilities tests. Academic Therapy, 1979, 14, 276-273.
- Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.
- Anderson, R. C. Schemata-directed processes in language comprehension. Technical report from the Center for the Study of Reading. University of Illinois at Urbana-Champaign, 1979, No. 50.
- Andre, T. Does answering higher-level questions while reading facilitate productive learning. Review of Educational Research, 1979, 49, 280-318.
- Banks, C., & Sinha, U. An item analysis of the Progressive Matrices test. British Journal of Psychology, 1951, 4, 91-94.
- Bloom, L., & Lahey, M. Language development and language disorders. New York: John Wiley and Sons, 1978.
- Boning, R. A. Specific Skill Series (B-F). Baldwin: Barnell Loft, 1978.
- Bransford, J. D. Human cognition: Learning, understanding and remembering. Belmont: Wadsworth, 1979.
- Bransford, J. D., Barclay, J. R., & Franks, J. J. Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 1972, 3, 193-209.
- Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.
- Bransford, J. D., & Johnson, M. K. Considerations of some problems of comprehension. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.
- Burke, H. R. Raven's Progressive Matrices: A review and critical evaluation. Journal of Genetic Psychology, 1958, 93, 199-228.
- Bush, C. L., & Huebner, M. H. Strategies for reading in the elementary school (2nd ed.). New York: Macmillan, 1979.

- Canney, G., & Winograd, P. Schemata for reading and reading comprehension performance. Technical report from the Center for the Study of Reading. University of Illinois at Urbana-Champaign, 1979, No. 120.
- Carroll, J. B. Psychometric tests as cognitive tasks: A new "Structure of Intellect". In L. B. Resnick (Ed.), The nature of intelligence. Hillsdale: Lawrence Erlbaum, 1976.
- Carroll, J. B. Remarks on Sternberg's "factor theories of intelligence are all right almost". Educational Researcher, 1980, 9, 14-18.
- Carroll, J. B., & Walton, M. Has the reel reading problem bin lade bear? Summary comments on the theory and practice of early reading. In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading Vol. III. Hillsdale: Lawrence Erlbaum, 1979.
- Cattell, R. B. Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 1963, 54, 1-22.
- Cattell, R. B. Abilities: Their structure, growth, and action. Boston: Houghton Mifflin, 1971.
- Chafe, W. Meaning and the structure of language. Chicago: University of Chicago Press, 1970.
- Chafe, W. Creativity in verbalization and its implications for the nature of stored knowledge. In R. O. Freedle (Ed.), Discourse production and comprehension. Norwood: Ablex Publishing Corporation, 1977.
- Chomsky, N. Syntactic structures. The Hague: Mouton, 1957.
- Cramer, E. H., Pascarella, E. T., & Hendress, C. An analysis of analogical reasoning and reading comprehension ability in middle and upper grade students. In M. L. Kamil & A. J. Moe (Eds.), Perspectives on reading research and instruction. Washington, D.C.: National Reading Conference, 1980.
- Cronbach, L. J. The two disciplines of scientific psychology. American Psychologists, 1957, 12, 671-684.
- Daneman, M., & Carpenter, P. A. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 450-466.
- Davidson, R. E. The role of metaphor and analogy in learning. In J. R. Levin & V. L. Allen (Eds.), Cognitive learning in children: Theories and strategies. New York: Academic Press, 1976.
- Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1978, 14, 481-533.
- Elley, W. B., & Macarthur, R. S. The Standard Progressive Matrices as a culture-reduced measure of general intellectual ability. Alberta Journal of Educational Research, 1962, 8, 54-65.

- Fillmore, C. J. Ideal reader - Real readers. A paper presented at the meeting of the Georgetown University Round Table, Washington, D.C., 1981.
- Frederiksen, C. H. Discourse comprehension and early reading. In L. B. Resnick & P. A. Weaver (Eds.), Theory and practices of early reading Vol. I. Hillsdale: Lawrence Erlbaum, 1979.
- Foulds, G. A., & Raven, J. C. Normal changes in the mental abilities of adults as age advances. Journal of Mental Science, 1948, 94, 133-142.
- Gabriel, K. R. The simplex structure of the Progressive Matrices test. British Journal of Statistical Psychology, 1954, 7, 9-14.
- Goldman, S. R., & Bisanz, J. Understanding the development of analogical reasoning ability. A paper presented to the annual meeting of the American Educational Research Association, Boston, April 1980.
- Goodman, K. S. Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 1967, 6, 126-135.
- Guilford, J. P. The three faces of intellect. American Psychologist, 1959, 14, 469-479.
- Guthrie, J. T., & Tyler, J. S. Psycholinguistic processing in reading and listening among good and poor readers. Journal of Reading Behavior, 1976, 8, 415-426.
- Hansen, J., & Pearson, P. D. The effects of inference training and practice on young children's comprehension. Technical report from the Center for the Study of Reading. University of Illinois at Urbana-Champaign, 1980, No. 166.
- Hayes, D. A., & Tierney, R. J. Increasing background knowledge through analogy: Its effects upon comprehension and learning. Technical report from the Center for the Study of Reading. University of Illinois at Urbana-Champaign, 1980, No. 186.
- Holden, C. A new visibility for gifted children. Science, 1980, 210, 879-880, 882.
- Horn, J. L. Organization of abilities and the development of intelligence. Psychological Review, 1968, 75, 242-259.
- Hornowski, B. Badanie zdolnosci myslenia testem percepcyjnym "Matrix" J. C. Raven. Filozofia, Psychologia, Pedagogika, 1959, 3, 17-33.
- Huck, S. W., Cormier, W. H., & Bounds, W. G. Reading strategies and research. New York: Harper & Row, 1974.
- Huttenlocher, J. The origins of language comprehension. In R. L. Solso (Ed.), Theories in cognitive psychology. Potomac: Lawrence Erlbaum, 1974.

- Huttenlocher, J. Language and intelligence. In L. B. Resnick (Ed.), The nature of intelligence. Hillsdale: Lawrence Erlbaum, 1976.
- Johnson, M. K., Bransford, J. D., & Solomon, S. Memory for tactic implications of sentences. Journal of Experimental Psychology, 1973, 98, 203-205.
- Kalina, W. The influence of the internal validity of a test on its external validity. Przegląd Psychologiczny, 1967, 14, 68-76.
- Keir, G. The Progressive Matrices as applied to school children. British Journal of Psychology, 1949, 2, 140-150.
- King, W. H. The development of scientific concepts in children. British Journal of Educational Psychology, 1963, 33, 240-252.
- Kintsch, W., & Kozminsky, E. Summarizing stories after reading and listening. Journal of Educational Psychology, 1977, 69, 491-499.
- Loranger, A. W., & Misiak, H. Critical flicker frequency and some intellectual functions in old age. Journal of Gerontology, 1959, 14, 323-327.
- Macarthur, R. S. Assessing the intellectual ability of Indian and Metis pupils at Fort Simpson, N. W. T. Ottawa: Department of Northern Affairs and National Resources, 1962.
- Macarthur, R. S. Mackenzie District norming project. Ottawa: Department of Northern Affairs and National Resources, 1968.
- Macnamara, J. On the relation between language learning and thought. In J. Macnamara (Ed.) Language learning and thought. New York: Academic Press, 1977.
- Miller, J. R., & Kintsch, W. Recall and readability of short prose passages. A paper presented at the meeting of the American Educational Research Association, Boston, April 1980.
- Moran, R. E. Progressive Matrices and the educationally disadvantaged. Mental Retardation, 1972, 10, 3-9.
- Nelson, K. Cognitive development and the acquisition of concepts. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale: Lawrence Erlbaum, 1977.
- Nisbet, J. D. Family environment and intelligence. Eugenics Review, 1953, 45, 31-40.
- Olson, D. R. Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 1970, 77, 257-273.
- Pearson, P. D., & Studt, A. Effects of word frequency and contextual richness on children's word identification abilities. Journal of Educational Psychology, 1975, 67, 89-95.

- Raven, J. C. The comparative assessment of intellectual ability. British Journal of Psychology, 1948, 39, 12-19.
- Raven, J. C. Standard Progressive Matrices Sets A, B, C, D and E. Cambridge: H. K. Lewis & Co., 1958.
- Resnick, L. B. Introduction: Changing conceptions of intelligence. In L. B. Resnick (Ed.), The nature of intelligence. Hillsdale: Lawrence Erlbaum, 1976.
- Rohwer, W. D., & Harris, W. J. Media effects on prose learning in two populations of children. Journal of Educational Psychology, 1975, 67, 651-657.
- Rohwer, W. D., & Matz, R. D. Improving aural comprehension in white and black children: Pictures versus print. Journal of Experimental Psychology, 1975, 19, 23-26.
- Rumelhart, D. E. Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale: Lawrence Erlbaum, 1980.
- Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale: Lawrence Erlbaum, 1977.
- Sachs, J. S. Memory in reading and listening to discourse. Memory and Cognition, 1974, 2, 95-100.
- Sheppard, C., Fiorentino, D., Collins, L., & Merlis, S. Raven's Progressive (1938): Normative data on male narcotic addicts. Psychological Reports, 1968, 23, 343-348.
- Shulman, L. S. Reconstruction of educational research. Review of Educational Research, 1970, 40, 371-396.
- Slosson, R. L. Slosson Intelligence Test (SIT) for Children and Adults. East Aurora: Slosson Educational Publications, 1977.
- Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. Journal of Educational Psychology, 1977, 69, 381-387.
- Smith, F. Psycholinguistics and reading. New York: Holt, Rinehart and Winston, 1973.
- Smith, F. Understanding reading (2nd ed.). New York: Holt, Rinehart and Winston, 1978.
- Smith, F. Conflicting approaches to reading research and instruction. In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading Vol. II. Hillsdale: Lawrence Erlbaum, 1979.

- Snow, R. E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-291.
- Snow, R. E. Toward a theory of aptitude for learning: I. Fluid and crystallized abilities and their correlates. Technical report from the Center for Advanced Study in the Behavioral Sciences. Office of Naval Research, 1979, No. N00014-75-C-0882.
- Snow, R. E. Aptitude processes. In R. E. Snow, P. A. Frederico, & W. E. Montague (Eds.), Aptitude, learning, and instruction Vol. I: Cognitive process analyses of aptitude. Hillsdale: Lawrence Erlbaum, 1980. (a)
- Snow, R. E. Intelligence for the year 2001. Intelligence, 1980, 4, 185-199. (b)
- Spearman, C. E. General intelligence, objectively determined and measured. American Journal of Psychology, 1904, 15, 201-292.
- Spearman, C. The nature of intelligence and the principles of cognition. New York: Macmillan, 1923.
- Spearman, C. The abilities of man. New York: Macmillan, 1927.
- Spiro, R. J. Etiology of reading comprehension style. In M. Kamil & A. Moe (Eds.), Reading research: Studies and applications. Washington, D. C.: National Reading Conference, 1979.
- Spiro, R. J. Constructive processes in prose comprehension and recall. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale: Lawrence Erlbaum, 1980.
- Spiro, R. J., & Tirre, W. C. Individual differences in schema utilization during discourse processing. Journal of Educational Psychology, 1980, 72, 204-208.
- Sternberg, R. L. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale: Lawrence Erlbaum, 1977.
- Sternberg, R. L. Factor theories of intelligence are all right almost. Educational Researcher, 1980, 9, 6-13, 18.
- Sternberg, R. L., & Rifkin, B. The development of analogical reasoning processes. Journal of Experimental Child Psychology, 1979, 27, 195-232.
- Sticht, T. G. Learning by listening. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: V. H. Winston, 1972.
- Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. Journal of Educational Psychology, 1917, 8, 323-332.

- Thurstone, L. L. Primary mental abilities. Psychometric Monographs, 1938, No. 1.
- Trabasso, T. On the making of inferences during reading and their assessment. Technical report from the Center for the Study of Reading. University of Illinois at Urbana-Champaign, 1980, No. 157.
- Tully, G. E. Test-retest reliability of the Raven Progressive Matrices Test (Form 1938) and the California Test of Mental Maturity, level 4 (S-F 1963). Florida Journal of Educational Research, 1967, 9, 67-74.
- Tyler, L. E. The intelligence we test - An evolving concept. In L. B. Resnick (Ed.), The nature of intelligence. Hillsdale: Lawrence Erlbaum, 1976.
- Vygotsky, L. V. Thought and language (E. Haufmann, & G. Vakar, Ed. and trans.). Cambridge: MIT Press, 1962.
- Whitely, S. E. Process outcome models for verbal aptitude. A paper presented at the meeting of the American Educational Research Association, Boston, April 1980.
- Whitely, S. E., and Schneider, L. M. Process outcome models for verbal aptitude. Technical report from National Institute of Education. University of Kansas, 1980, No. NIE-80-1.
- Wilkinson, A. C. Children's understanding in reading and listening. Journal of Educational Psychology, 1980, 72, 561-574.

APPENDIX

A

Data from Post Hoc

Pilot Study

Table A
Raw Score Data from Post Hoc Pilot Study

	Subjects	<u>SPM</u>	Silent Reading*	Listening*	Slosson
Boys <u>Third</u> Graders	1	48	50	57	187
	2	21	10	13	106
	3	19	3	20	140
	4	31	14	22	130
	5	32	14	12	108
	6	24	6	11	97
Boys <u>Fifth</u> Graders	1	53	58	60	164
	2	46	35	58	122
	3	27	30	13	98
	4	40	56	55	112
	5	31	6	21	87
	6	23	32	13	92

*Data based on inventories of 60 items each.

Table B

Means and Standard Deviations from Post Hoc Pilot Study

	SPM		Silent Reading		Listening		Slosson	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Boys								
Third Grade	29.33	13.22	21.00	20.70	30.00	19.30	144.33	33.21
Girls								
Total	29.17	9.69	16.17	15.65	22.50	15.96	128.00	30.20
Boys								
Fifth Grade	42.00	10.98	41.00	12.19	43.67	21.70	128.00	27.28
Girls								
Total	36.67	10.62	36.17	17.50	36.67	21.22	112.50	25.90

Table C

Correlation Coefficients from Post Hoc Pilot Study

<u>SPM</u>	Silent Reading			
	.7990			
		Listening		
	<u>SPM</u>	.9131		
			Slosson	
		<u>SPM</u>	.5880	
				Silent Reading
			Listening	.8175

APPENDIX

B

Data from Procedural Pilot

Table D
Raw Score Data
from Procedural Pilot

	Subjects	SPM	Silent Reading*	Listening*	Slosson
Boys <u>Third</u> <u>Graders</u>	1	10	5	3	80
	2	14	5	5	102
	3	18	6	6	111
	4	11	2	4	79
Girls	5	35	2	6	100
	6	26	6	6	106
Boys <u>Fifth</u> <u>Graders</u>	1	22	21	22	106
	2	27	12	15	109
	3	20	19	25	96
	4	14	3	5	96
Girls	5	30	14	12	108
	6	37	20	15	108

*Data based on inventories of 60 items each.

Table E

Means and Standard Deviations for Data from Procedural Pilot

	SPM		Silent Reading		Listening		Slosson	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Boys								
Third Grade	19	9.76	4.3	1.86	5	1.26	96.33	13.57
Girls	13.25	3.59	4.5	1.73	4.5	1.30	93	16.02
Total	30.50	6.36	4	2.83	6	0	103	4.24
Boys								
Fifth Grade	25.00	8.10	14.83	6.80	15.67	7.15	103.83	6.14
Girls	20.75	5.38	13.75	8.14	16.75	8.88	101.75	6.75
Total	33.50	4.95	17.00	4.24	13.50	2.12	108	0

Table F

Correlation Coefficients from Procedural Pilot

<u>SPM</u>	Silent Reading			
	.432			
		Listening		
	<u>SPM</u>	.356		
			Slosson	
		<u>SPM</u>	.652	
				Silent Reading
			Listening	.919

APPENDIX

C

Scoring Materials

COMPREHENSION PREDICTION INVENTORY I

SCORING CODES

S = SILENT READING

L = LISTENING

✓ = ACCURATE RESPONSE

X = INACCURATE RESPONSE

LEVEL B ()

Ex. B6-1 (3) _____

B7-4 (1) _____

B8-3 (2) _____

B9-1 (2) _____

B10-2 (3) _____

B11-3 (2) _____

_____ %

LEVEL C ()

Ex. C8-2 (1) _____

C9-4 (2) _____

C10-5 (1) _____

C11-1 (1) _____

C12-4 (3) _____

C13-1 (3) _____

_____ %

INVENTORY I

LEVEL D ()

 Ex. D12-3 (3) _____

D13-1 (3) _____

D14-5 (2) _____

D15-4 (3) _____

D16-1 (3) _____

D17-4 (2) _____

 _____%

LEVEL F ()

 Ex. F14-3 (4) _____

F15-1 (1) _____

F16-3 (3) _____

F17-2 (4) _____

F18-4 (3) _____

F19-5 (2) _____

 _____%

LEVEL E ()

 Ex. E4-1 (2) _____

E5-2 (2) _____

E6-4 (4) _____

E7-2 (1) _____

E8-1 (4) _____

E9-5 (2) _____

 _____%

LEVEL G ()

 Ex. G11-4 (4) _____

G12-2 (3) _____

G13-5 (3) _____

G14-1 (3) _____

G15-3 (3) _____

G16-2 (3) _____

 _____%

COMPREHENSION PREDICTION INVENTORY II

SCORING CODES

S = SILENT READING

L = LISTENING

✓ = ACCURATE RESPONSE

X = INACCURATE RESPONSE

LEVEL B ()

Ex. B8-2 (1) _____

B12-4 (2) _____

B13-2 (2) _____

B14-3 (2) _____

B15-4 (2) _____

B16-2 (3) _____

_____ %

LEVEL C ()

Ex. C13-4 (3) _____

C14-3 (2) _____

C15-2 (2) _____

C16-2 (1) _____

C17-1 (3) _____

C18-3 (2) _____

_____ %

INVENTORY II

LEVEL D ()

Ex. D10-1 (2) _____

D18-2 (3) _____

D19-2 (3) _____

D20-5 (3) _____

D21-2 (2) _____

D22-3 (3) _____

_____ %

LEVEL F ()

Ex. F25-1 (2) _____

F20-2 (2) _____

F21-3 (4) _____

F22-5 (1) _____

F23-4 (3) _____

F24-2 (3) _____

_____ %

LEVEL E ()

Ex. E1-1 (3) _____

E10-1 (4) _____

E11-1 (3) _____

E12-5 (3) _____

E13-1 (4) _____

E14-4 (3) _____

_____ %

LEVEL G ()

Ex. G6-3 (2) _____

G17-5 (1) _____

G18-4 (2) _____

G19-2 (2) _____

G20-5 (3) _____

G21-1 (4) _____

_____ %

<u>LEVEL D</u>	
Ex. D12-3	_____

D13-1	_____
D14-5	_____
D15-4	_____
D16-1	_____
D17-4	_____

Response Card for Set D
Inventory I

APPENDIX

D

Sample Linguistic and
Non-Linguistic Items

C13-1

THE SOUND OF THIS BEETLE:

- A. PLEASED PEOPLE
- B. SURPRISED PEOPLE
- C. FRIGHTENED PEOPLE

Question and Response-Alternatives for Sample
Passage from Barnell-Loft Specific Skill Series
Drawing Conclusions Set C

F19-5

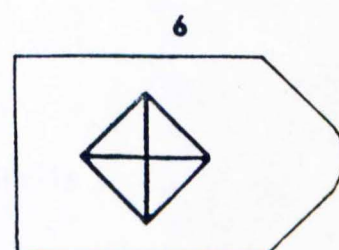
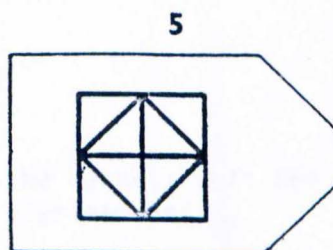
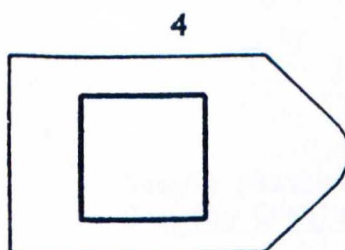
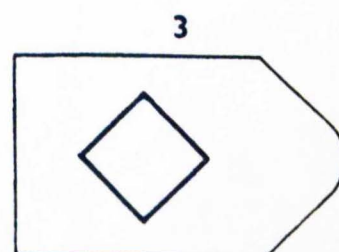
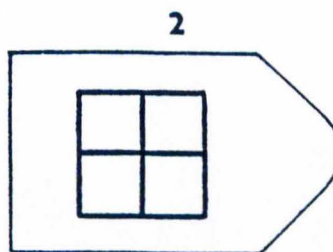
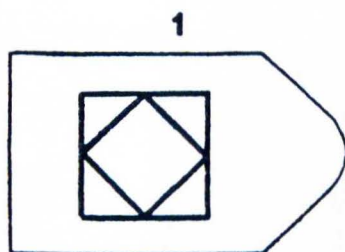
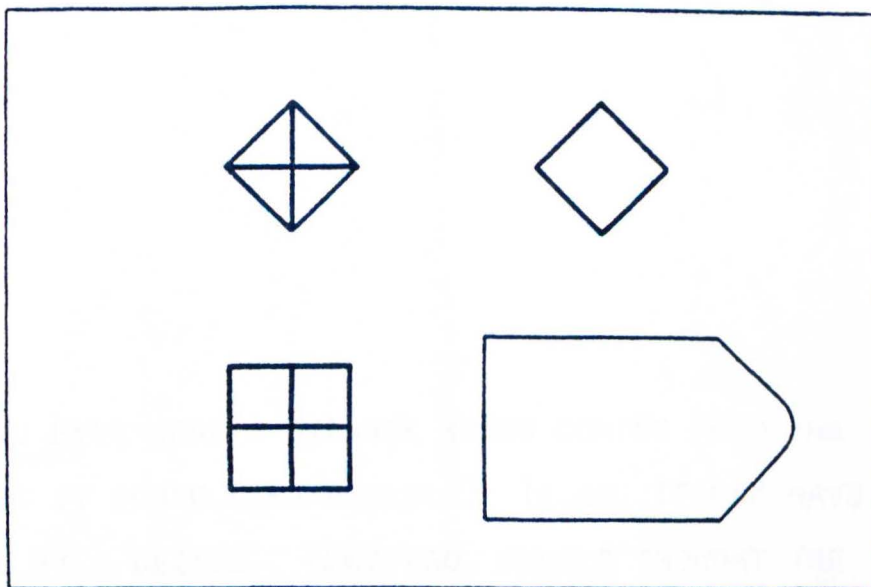
LACROSSE IS THE OLDEST ORGANIZED SPORT PLAYED IN NORTH AMERICA. THIS FAST, ROUGH GAME WAS PLAYED BY THE INDIANS. EARLY FRENCH AND ENGLISH SETTLERS, AFTER WATCHING INDIANS PLAY THE GAME, ADOPTED IT AND MADE SOME CHANGES IN THE RULES. PLAYERS ARMED WITH LONG STICKS TRY TO THROW OR KICK A BALL INTO THE OPPOSING TEAM'S GOAL AREA.

F19-5

FROM THE STORY YOU CAN TELL THAT:

- A. LACROSSE IS A VERY POPULAR GAME
- B. EARLY SETTLERS LEARNED FROM THE INDIANS
- C. NO CHANGES HAVE EVER BEEN MADE IN LACROSSE
- D. ALL PEOPLE PLAYED LACROSSE

Question and Response-Alternatives for Sample Passage
from the Barnell-Loft Specific Skill Series Drawing
Conclusions Set F

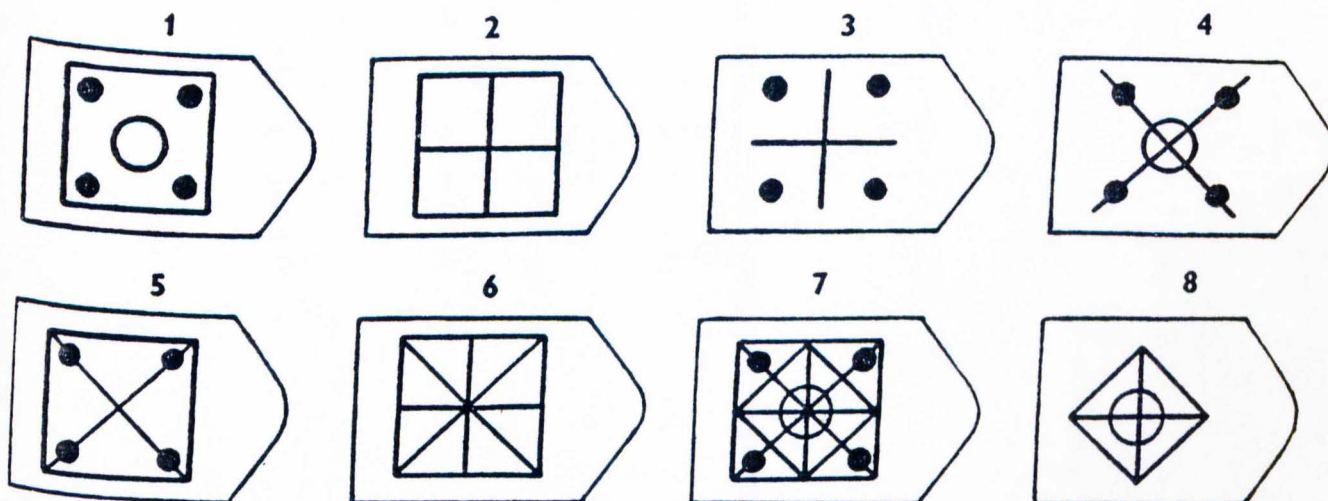
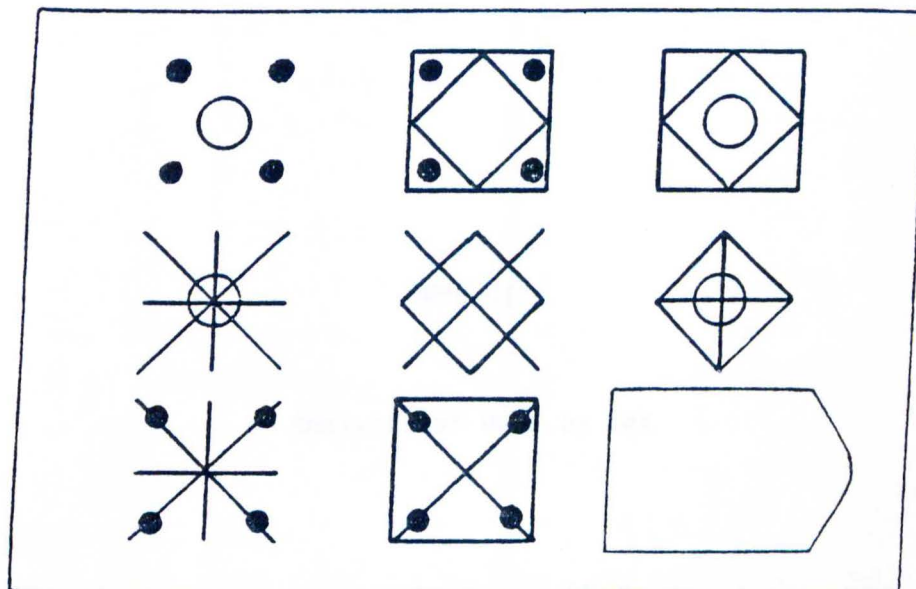
B II

Sample Item from the Raven's Standard Progressive
Matrices Set B

C13-1

DID YOU EVER HEAR A STRANGE SOUND COMING FROM THE WALL? DID IT SOUND LIKE A CLOCK? IF SO, IT MAY HAVE BEEN MADE BY A BEETLE. LONG AGO, PEOPLE THOUGHT THE TICKING MEANT THAT SOMEONE WAS ABOUT TO DIE. THUS THE BEETLE IS CALLED "THE DEATHWATCH BEETLE."

E 10



Sample Item from the Raven's Standard Progressive
Matrices Set E

APPENDIX

E

Analysis of Data by Sex

Table G
Means and Standard Deviations of
Outcome Measures by Sex

	SPM		Silent Reading		Listening		Slosson	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
<u>Boys</u>								
Third Grade	26.0	10.712	7.5	6.856	9.125	7.769	128.625	24.844
Fifth Grade	38.4	8.1756	18.8	8.28	16.7	8.866	118.9	21.566
Total	32.889	11.229	13.778	9.513	13.333	9.201	123.222	23.681
<u>Girls</u>								
Third Grade	30.167	8.009	7.583	6.982	16.917	29.823	113.167	12.089
Fifth Grade	37.7	8.741	15.0	6.115	18.3	8.718	111.3	14.022
Total	33.59	9.154	10.5	7.668	13.0	8.252	112.318	13.036
<u>t</u>	-212, $p>.05$		1.18, $p>.05$		1.17, $p>.05$		1.8, $p>.05$	

APPENDIX

F

Data from Study

TABLE H
Data for Third-Grade Subjects for Study

Subjects	Raven's	Silent Reading	Listening	Slosson
1	30	4	15	119
2	18	3	9	104
3	13	5	9	118
4	26	5	4	102
5	34	5	5	130
6	44	27	14	115
7	40	5	5	102
8	30	9	5	122
9	22	5	9	124
10	22	3	5	115
11	40	15	14	125
12	19	5	5	121
13	48	25	29	187
14	31	5	9	130
15	32	5	4	108
16	14	2	8	140
17	21	4	4	106
18	24	3	4	97
19	32	13	14	130
20	30	3	5	92

TABLE I
Data for Fifth-Grade Subjects for Study

Subjects	Raven's	Silent Reading	Listening	Slosson
1	40	28	15	133
2	50	19	10	127
3	25	4	5	80
4	33	10	19	99
5	33	28	28	134
6	43	15	30	117
7	39	14	14	116
8	45	14	15	106
9	32	19	15	129
10	50	14	24	121
11	30	14	29	123
12	35	15	14	104
13	44	28	14	119
14	42	14	14	119
15	53	29	30	164
16	46	15	29	122
17	40	28	28	112
18	31	3	9	87
19	23	14	4	92
20	27	13	4	98