THE DEVELOPMENT OF A SINGLE FACTOR LISTENING COMPREHENSION TEST

By

PEGGY ALEXANDER KERR

Bachelor of Arts
Central State University
Edmond, Oklahoma
1962

Master of Education
Central State University
Edmond, Oklahoma
1976

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
MAY, 1989

THE DEVELOPMENT OF A SINGLE FACTOR

LISTENING COMPREHENSION TEST

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGMENTS

This study has been successfully completed only with the help and support of many special people. First, I would like to express my sincere appreciation to my committee for your suggestions and guidance. A special thanks goes to Dr. Paul Warden, my committee chairman and dissertation advisor, for your support and faith in my ability to accomplish this task.

This study required the involvement of so many people. To the children who took the test, parents who gave your permission, school administrators who allowed me to come into your schools and teachers who allowed me to come into your classrooms to test, thank you all. I would especially like to express my appreciation to my students who spent many hours administering the test during the norming stage of the study.

The last and best thank you goes to my husband, Michael. You have been patient, kind, and understanding throughout this effort. Not only have you given unfailing emotional support, but you have shared your knowledge of test construction and provided invaluable suggestions along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Statement of the Problem

Literacy is composed of four skills--listening, speaking, reading, and writing. These skills have a hierarchical interdependence. Lundsteen (1971) observed, "Chronologically, children listen before they speak, speak before they read, and read before they write" (p. 1). Reading, writing, and speaking have long been integral parts of the educational process. Yet, while listening is the first language skill to appear, it receives the least attention in education. Research, curriculum development and assessment procedures in listening have lagged far behind developments in the other skill areas of literacy. While most educators and researchers would readily agree that listening is essential to learning, the absence of a test which directly measures listening skill has impeded the investigation of its importance.

Two recent public education mandates have had an impact on listening research. The first of these was the Education For All Handicapped Children Act of 1975 (P.L. 94-142) which required a measure of listening comprehension for every child who was evaluated for a learning disability.

The second mandate was The Education Amendments of 1978 (P.L. 95-561) which established legal definitions of basic skills in education. In the area of "effective oral communication," speaking and listening were identified as the basic skills in that process. As a result of these mandates, many states are actively seeking appropriate listening assessment instruments. They are finding existing ones to be unsatisfactory (Brown, Backlund, Gurry, and Jandt, 1979; Faires, 1980; Rubin, Daly, McCroskey and Mead, 1982).

The State Department of Education in Oklahoma selected the Listening Comprehension Subtest of the Stanford Achievement Test, to fulfill the P.L. 94-142 requirement after failing to find a satisfactory instrument. This subtest was designed to be administered as part of the larger Stanford Achievement Test battery and to be given in a group setting. In Oklahoma, only the single subtest is administered to individual children. Also, although the subtest is administered individually, the examinee is given a booklet containing printed options for the questions. This may give the examinee who reads well an advantage over one who reads poorly. Therefore, the validity and reliability of this subtest comes into question because Oklahoma has departed from standardization procedures in both administration and scoring.

Every state needs a standardized listening comprehension test to fulfill the P.L. 94-142 requirement.

However, a reliable and valid instrument which is designed to be administered individually is currently unavailable. The need for such an instrument is pressing.

Purpose of the Study

The main purpose of this study was to develop a listening comprehension test, standardize the instrument and determine its reliability and validity. A secondary purpose of this study was to determine whether there were any significant differences between populations of referred and non-referred students on listening comprehension.

Definition of Terms

Listening comprehension is defined as the process of deriving meaning from spoken symbols. For the purposes of this study, listening comprehension is operationally defined as the amount of information correctly recalled by the listener (in the form of a score on the test) after having been given a verbal presentation of passages followed by questions.

Reading comprehension is constitutively defined as the process of deriving meaning from printed symbols.

Non-referred is defined as a descriptor of a child who has never been referred for an individual psycho-educational or psychological evaluation.

Referred is defined as a descriptor of a child who has been referred, at some time, for an individual psycho-

educational evaluation and subsequently was placed in a learning disabilities class. This does not include children who have been referred for gifted and talented evaluation or children who have been identified as mentally retarded.

Background

While the literature on listening variously defines the construct of listening comprehension, there is little or no agreement among theorists. This is one of the primary reasons for the paucity of listening comprehension tests. None of the definitions which have been offered are empirically based (Binford, 1977; Devine, 1978). The definitions include logical descriptions of the listening process (Lundsteen, 1971) and lists of basic competencies representing a series of listening subskills (Mead, 1978; Devine, 1981). Listening has been defined by theorists in foreign language acquisition (Rixon, 1981), linguistics (Pearson and Fielding, 1983), psycholinguistics (Clark and Clark, 1977), and sociolinguistics (Hymes, 1964) in terms of underlying cognitive processes the listener uses to comprehend a spoken message. These various theorists view listening comprehension in ways unique to their disciplines, and their definitions reflect their differences in orientation.

An operational definition which states the observable and measurable behaviors of listening comprehension can serve as a conceptual base for a listening comprehension

test. The test can then provide information that clarifies the construct of listening comprehension.

Because some listening comprehension tests were designed to predict reading comprehension potential (Duker, 1971), a number of studies (Sticht, Beck, Hauke, Kleiman and James, 1974) have examined the relationship between listening comprehension and reading comprehension. In reviewing these studies, Schell (1981) concluded that measures of listening comprehension cannot be valid predictors of reading comprehension. According to P.L. 95-561, listening comprehension must now be considered important in its own right. A test designed to exclusively measure listening comprehension is clearly needed.

In the search for elements unique to listening comprehension or reading comprehension, theorists (Devine, 1978; Lundsteen, 1971) have compared the activities of listening and reading and found them to be different in several ways. Primarily, a listener has the situational advantage of gaining information from the prosodic (visual, stress, pitch, and juncture cues) features of the auditory message but, on the other hand, needs to remember what was said. The reader has the advantages of stopping, looking ahead, and rereading. These differences are the very elements which should distinguish reading comprehension and listening comprehension tests. However, the existing tests for both skills are more alike than different.

A great deal of the listening research has been directed toward drawing comparisons between listening comprehension and reading comprehension. Researchers (Brown, 1965; Faires, 1980) have used correlational studies to investigate whether tests of listening comprehension and reading comprehension were measuring the same underlying mental processes. The results of these studies have yielded a wide range of coefficients, giving rise to questions concerning the reliability and validity of listening comprehension tests (Devine, 1978).

Published lists of curriculum objectives for listening comprehension have been found to correspond with objectives for reading comprehension. Apparently, the authors of those lists viewed listening and reading as being so alike that listening comprehension and reading comprehension could be broken down into similar, specific skills. This conclusion may have led to the assumption that tests suitable for reading comprehension might also be suitable for listening comprehension (Brown, Backlund, Gurry and Jandt, 1979). Tests of the two skills have generally paralleled each other in their construction, skills measured and types of questions asked (Binford, 1977). The tests typically use multiple choice questions which require reading and recognition rather than recall, thus contaminating listening with reading skills (Faires, 1980).

The listening literature has also addressed the relationship between listening comprehension and

intelligence. While intending to show that listening comprehension could be accounted for by reasoning and verbal comprehension, Spearritt (1961) isolated a listening comprehension factor. Mead (1978) examined the correlation between listening comprehension and verbal intelligence, reporting coefficients of .59 and .47. Devine (1978) concluded that the observed relationships between listening comprehension and intelligence are small enough to indicate that listening depends upon factors other than intelligence. Moreover, questions must be raised about the validity of the listening tests used to establish these correlations.

Many theorists have been critical of listening comprehension tests due to lack of agreement on the underlying construct of listening (Faires, 1980; Rubin, Daly, McCroskey and Mead, 1982; Watson and Barker, 1985). These criticisms have spurred recommendations for listening comprehension test construction (Bostrom and Waldhart, 1980; Backlund, Brown, Gurry and Jandt, 1982; Mead and Rubin, 1985). Many of the recommendations reflect the unique viewpoints of the authors' different disciplines. For example, theorists have specifically disagreed on whether listening is a series of subskills or a unitary skill, and whether a listening comprehension test should attempt to measure these subskills or yield a single raw score (Binford, 1977).

A number of researchers have explored variables which seem to influence listening comprehension and may therefore

be useful in developing instructional strategies for the teaching of listening comprehension (Woodcock and Clark, 1968; Schunk and Rice, 1983; Townsend, 1983). A review of these studies provides some indications that listening skills could be improved through teaching strategies (Pearson and Fielding, 1983). However, research in curriculum development has been slowed by a lack of listening comprehension tests with demonstrated reliability and validity.

Methodology

The listening comprehension test developed in this study was constructed to avoid the shortcomings of previous instruments and was rigorously tested for meaningful levels of reliability and validity. The test has one form across grade levels K-6 and ages 5-12. It contains five stimulus passages and 50 questions which are administered verbally and require a verbal or written response from the examinee.

The samples included children at grade levels K-6, both male and female, and either referred-for-testing or non-referred. The sample sizes were large enough to ensure the stability of obtained reliability and validity coefficients: $n > 100$ for each of two field test samples, $n > 300$ for the non-referred (standardization) sample, and $n > 50$ for the referred sample.

Two forms of the listening comprehension test were field tested. The test items were then quantitatively

evaluated and combined into the final form. A scoring guide was developed from the field test responses.

Three kinds of reliability were investigated--internal consistency, test-retest, and inter-examiner. Internal consistency reliability was estimated by computing a K-R 20 coefficient. Test-retest reliability was investigated by testing the same group one week apart and correlating their test and retest scores. Inter-examiner reliability was investigated by asking alternate examiners to score the same tests, using audiotapes of test administrations, and correlating their assigned scores.

Three kinds of validity were investigated--content, concurrent, and factorial. Content validity was built into the test by selecting passages from a larger sample of material which was written according to four pre-selected criteria. Concurrent validity was investigated by determining the correlation between the test scores and listening in a classroom situation. Factorial validity was investigated through factor analysis of the non-referred sample scores.

Hypotheses

1. Given the methods and procedures described in this study, a reliable and valid test of listening comprehension will be constructed.

2. There will be significant differences between referred and non-referred children on listening comprehension.

Implications

A valid test of listening comprehension which is standardized for individual administration and uncontaminated by reading requirements is currently unavailable. If a valid and reliable test of listening comprehension can be developed, accurate diagnosis and effective remediation of learning difficulties related to listening comprehension can occur.

CHAPTER II

REVIEW OF THE LITERATURE

Historically, research in the area of listening has involved three major areas. This chapter will deal with each of these areas, which are as follows: (a) defining listening, (b) relationship between listening comprehension and other factors, and (c) recommendations for development of listening comprehension tests. There is also a body of research on instructional implications for the teaching of listening comprehension. Although that literature is of little importance to the present study, a brief review of major findings is included.

Defining Listening

Brown (1954) preferred the term, "auding" to "listening comprehension." He reasoned that listening was not confined to language. For example, one could hear and interpret other sounds, such as a dog barking, and experience listening comprehension. Therefore, auding was defined "as the gross process of looking at, recognizing, and interpreting spoken symbols" (p. 128). Although auding is used synonymously for listening and listening comprehension, it is not a commonly used term in the literature.

Binford (1977) reviewed studies conducted from 1926 through 1969 and extracted 18 separate definitions of listening comprehension. Binford concluded that the authors of the definitions deserve recognition for their contributions, but the definitions were "never conceptually modified or refined beyond their initial use" (p. 17). These definitions were not empirically-based. The authors simply offered logical descriptions of listening.

Devine (1978) summarized the listening research of the previous 50 years. Among the definitions he found, his favorite was Lundsteen's (1971), which states that listening comprehension is "the process by which spoken language is converted to meaning in the mind" (p. 1). Devine commented:

> This definition is not the result of research, in the sense that some of our definitions of 'learning' or 'intelligence' rest on theory supported by controlled, experimental studies. It is, rather, the result of careful analysis of what seems to be known to date. (p. 297)

Neither Binford (1977) nor Devine were able to find the construct of listening empirically defined in the literature.

Listening comprehension has also been defined in terms of basic competencies by those involved in listening instruction and measurement. These lists usually form the basis for a test. Mead (1978) reviewed a pilot listening assessment initiated by the National Assessment of

Educational Progress (NAEP) and the Speech Communication Association (SCA) in which educators and content specialists listed competencies for listening comprehension which they considered most important. Then test items were designed to measure these competencies. Test items required students to recall significant details, comprehend main ideas, interpret and draw inferences, and make judgments about the speaker and information presented. Mead pointed out that these objectives were based on logic and not on empirical evidence. Apparently test items were developed without a working definition of listening comprehension.

In 1981, Devine developed the Listening Skills Assessment for the New Hampshire State Department of Education. He reported that this instrument was "developed on the assumption that listening is best viewed--at least for teaching and testing purposes as a composite of discrete skills" (p. 5). In order to establish content validity, a search of the listening literature was conducted which resulted in a master list of 40 listening skills. He classified these under five headings--simple recall, recognizing and following spoken directions, recognizing a speaker's purpose and plan, critical listening, and higher-level listening skills. The 45 multiple choice test items on the full form measured 53 specific listening skills in these five categories.

Both the NAEP/SCA and Listening Skills Assessment instruments defined listening, not as a unitary skill, but

as a series of subskills. These subskills evolved from "expert" opinion, not research.

Dunkel (1986), in discussing the importance of listening comprehension development at the initial stages of foreign language acquisition, reviewed Rixon's (1981) hypothesis concerning the steps an efficient listener goes through in extracting meaning from spoken symbols. These included finding a reason for listening, predicting what will be heard, deciding what to ignore, and checking one's understanding of the message in a variety of ways (e.g., by asking or answering questions, by carrying out specific tasks, or by making appropriate paralinguistic responses). Rixon stated that "the listener may jump backward or forward through steps as necessary" (p. 100). Rixon's description of listening is clearly offered as a hypothesis and not as a result of research.

Pearson and Fielding (1983) gave a linguistic analysis of listening comprehension. They stated that "listening comprehension seems to be controlled by the same set of cognitive processes as reading comprehension" (p. 1). Their definition reflected this comparison. They indicated that when a listener can orchestrate phonology (sound structure), syntax (sentence structure), semantics (word meanings and the relationships among meanings), and text structure (conventions about how events and assertions in narratives are typically structured) and apply them to achieve

interpretation of a text, he has experienced listening comprehension.

Landolfi (1984) discussed the process of listening comprehension from the psycholinguist viewpoint. She reviewed Clark and Clark's (1977) four step analysis of the comprehension process. In summary, the listener takes in raw speech, organizes it into constituents, constructs propositions and groups them together to form a message, and holds these in long-term memory while the original message is deleted. Landolfi also included a sociolinguistic viewpoint on listening by Hymes (1964) which states that "the listener has to consider who speaks to whom, why, where, and when if the interpretation of the message is to be adequate" (p. 8).

The definitions offered by these theorists are based on their "analyses" of the listening process given the perspective of their particular disciplines. These definitions include elements that go beyond observable behavior. They are broader and more refined than logical definitions but they have not been empirically validated.

Roberts (1985) examined five listening assessment tests--focusing on the strengths, weaknesses, procedural problems, and conceptualizations of each. He concluded that a definition of listening had not been agreed upon by those interested in listening research.

In summary, four points can be discerned:

1. The terms, auding, listening and listening

comprehension, are used synonymously by most theorists.

2. Listening has been described both as a unitary skill and as a series of subskills.

3. There are no empirically-based definitions of listening.

4. There is no agreement among theorist upon a definition of listening. Theorists in linguistics, psycholinguistics, sociolinguistics, and foreign language acquisition disagree on the cognitive processes the listener uses to comprehend the spoken message because they view listening from quite different perspectives.

The various definitions of listening comprehension have two elements in common: (a) The listener hears spoken language, and (b) the listener derives meaning from it. These elements appeared in the logical definitions offered by Lundsteen (1971) and others. In the present study, the definition of the construct of listening comprehension includes only these two common elements. This study does not attempt to define listening comprehension by describing an underlying process which is not observable. Tests require operational definitions which state observable and measurable behaviors. Such a definition is given in this study.

Relationship Between Listening Comprehension and Other Factors

Listening and Reading

The first test of listening was prepared by Donald D. Durrell and Helen B. Sullivan of Boston University in the 1930's. The test was named the Durrell Sullivan Reading Capacity Test (Duker, 1971). Apparently these authors believed that a measure of listening could predict reading aptitude. Investigations into this relationship have continued to the present.

Schell (1981) examined data from numerous studies investigating the relationship between listening comprehension and reading comprehension including a review of 31 such studies by Sticht, Beck, Hauke, Kleiman, and James (1974). Schell found that in grades one through six, almost all of the comparisons favored the listening comprehension mode. Schell concluded that reading comprehension and listening comprehension do not become approximately equal in achievement until about grade six; and reading comprehension grows far faster than listening comprehension, starting far below it but eventually catching up and surpassing it. Schell recommended that educators not accept listening comprehension measures as valid substitutes for measures of academic aptitude in determining reading potential.

In an effort to separate those elements which are unique to reading and listening, Devine (1978) delineated their linguistic differences. Concerning situational contexts, Devine pointed out that the reader is alone with the printed material; but the listener may gain information apart from what is being said. For example, the listener can observe the speaker's body language, stress, pitch, and juncture patterns. (Linguists refer to these patterns as "prosodic" features of an auditory message.) Concerning time context, the reader may stop, look ahead and reread the printed material; but the listener must remember what was said.

Lundsteen (1971) also compared listening and reading by listing the attributes for each activity. She included the element of emotion as potentially being quite strong in listening because listening may involve interaction with the speaker. She indicated that emotion is usually weak in reading.

While Devine (1978) conceded obvious situational and time context differences between listening and reading, he also offered the hypothesis that they were alike in three ways. First, he stated that "both are concerned with the intake half of the communications process" (p. 301). Devine did not explain his point, but he was probably referring to the idea that the four skills of literacy can be separated into intake (reading and listening) and output (writing and speaking).

Secondly, Devine (1978) stated that "the mental processes underlying both listening and reading must be at least somewhat similar" (p. 301). He reported correlations between test scores in reading and listening obtained from the same populations as being high and positive, but also acknowledged that others have raised questions about the validity of the tests used to establish these correlations.

The literature is replete with reports of correlations between listening tests and reading assessments used in their validation process or subsequent research studies. Brown (1965) reported correlations ranging from .76 to .82 between listening and reading for fourth, fifth, and sixth graders. Faires (1980) reported ten such coefficients ranging from .31 for high school and college students to .68. for kindergarten and first grade students.

Devine's (1978) third point stated that listening and reading are alike because "each seems to be a complex of related skills components" (p. 302). He reported that both reading and listening have been broken down into similar specific skills in the literature. This point was also addressed by Binford (1977) and Brown, Backlund, Gurry and Jandt (1979).

In a review of published listening programs, circulated teaching materials on listening and public schools' curriculum objectives on listening, Binford (1977) extracted 26 listening comprehension objectives, many of which closely paralleled reading.

When Brown et al. (1979) reviewed published lists of basic skills in reading and listening, they noticed a strong relationship between them. They found that several reading skills, such as recognizing and identifying meanings of words and phrases, identifying and understanding main ideas, and following directions, were common to both lists. Brown et al. indicated that this perceived relationship must be treated with caution because it could lead to the faulty assumption that assessment instruments suitable for one mode might also be suitable for another mode. While such a conclusion is unfounded, it is not uncommon. For example, after Binford (1977) reviewed 83 tests of listening, he stated, "All the tests were constructed on the basis of judgmental or logical analysis as perceived by authorities in the field. They were generally found to parallel reading tests in construction, skills measured, and types of questions asked" (p. 65).

Typically, listening tests involve presentation of a passage (by tape or examiner) followed by multiple choice questions. This method is used because most listening tests are designed for group presentation and requiring a verbal response would not be feasible. There are two problems with this design. First, the skill of recognition, rather than recall, is being measured. One of the primary ways in which reading and listening are different is that the listener must rely on his memory to recall what has been heard (Devine, 1978). Second, the multiple choice method requires

that the examinee read. Reading as a component of a listening test confounds test results. This point was addressed by Faires (1980) who stated, "Many assessments of listening ability are measured with a paper and pencil test and require some reading ability which contaminates listening results " (p. 17).

In summary, listening and reading have been compared to determine if listening ability could predict reading ability. There is considerable agreement among researchers that listening is not a good predictor of reading because they do not grow at equal rates and nearly all primary grade children comprehend better by listening than by reading.

Listening and reading have been compared in regard to their situational and time context differences. Devine (1978) delineated the unique advantages accruing to each mode. Only the listener can gain additional information through the prosodic features of the message and only the reader may stop, look ahead and reread. Because of the interpersonal potential of listening, Lundsteen (1979) included the element of emotion in her comparison of listening and reading.

Researchers have compared the underlying mental processes of listening and reading by examining the correlations between them. However, as Devine (1978) acknowledged, listening studies are of limited value unless the instruments are reliable and valid measures.

Listening and reading have also been compared in terms of their skills components. Published lists of basic skills in reading and listening seem to be based on an assumed commonality between them and have led to listening comprehension and reading comprehension tests which are parallel in construction, especially in the use of multiple choice questions.

Innovative research is needed to separate the elements unique to reading and listening. Such research is dependent on valid and reliable instruments. In the present study, listening comprehension is measured through recall and without any reading on the part of the examinee, thus minimizing the parallel to reading test construction.

Listening and Intelligence

Intending to demonstrate that listening comprehension could be accounted for by reasoning and verbal comprehension, Spearritt (1961) was the first to analyze the factorial structure of listening comprehension tests. He employed correlations from 34 tests given to 300 sixth grade pupils between the ages of 11 and 12. The battery of tests measured inductive (three tests), deductive (two tests), and general reasoning (one test); reading comprehension (six tests); attention (three tests); meaningful (two tests), rote (two tests), and span memory (two tests); auditory ability (four tests); and listening comprehension (nine tests). Eight factors were identified as having significant

loadings: induction, deduction, memory span, rote memory, meaningful memory, auditory resistance, verbal comprehension and listening comprehension. The listening comprehension factor was clearly demonstrated to be confined within the tests of listening comprehension. The factor was positively correlated with verbal comprehension, memory span, and reasoning. It was not highly related to the factor of attention.

Devine (1978) reviewed the literature from 1926 to 1961 investigating the correlations between listening tests and tests of intelligence. Finding coefficients of correlation ranging from .22 to .78, he stated that "there is enough variance in scores on the two kinds of tests, not accounted for by the elements common to both, to conclude that listening does depend upon something besides intelligence" (p. 300).

Mead (1978) reported significant differences between the responses of minority and nonminority students on a pilot listening skills assessment conducted by the National Assessment of Educational Progress and the Speech Communication Association (NAEP/SCA). An advisory group, convened to analyze test results, hypothesized that such extraneous factors as vocabulary, length of speech, interest in the content presented, accent, delivery rate, or test environment may have contributed to the differences in performance. A follow-up study was conducted to alleviate the problems identified by the reviewers of the pilot

assessment. Materials from the earlier study were rewritten and steps were taken to remove possible problems of bias by controlling the vocabulary level, decreasing length, increasing interest and using better tape recordings.

The final set of listening stimuli and questions (two packages) and a 25-item vocabulary test, a subtest of the Lorge-Thorndike Intelligence Test (to estimate verbal ability) were administered to 170 eleventh graders in four parts of the United States. Students who took package one had a mean listening score of 59.5 and a mean vocabulary score of 11.77. Students who took package two had a mean listening score of 55.9 and mean vocabulary score of 10.66. Correlations between listening ability and verbal ability for the two packages were .59 and .47. The mean vocabulary score of the minority students responding to package one was 12.8 and the mean vocabulary score of the minority students responding to package two was 6.7. For package one, there were no significant correlations between listening ability and minority status nor between verbal ability and minority status. For package two, the correlation between listening ability and minority status was -.23 and between verbal ability and minority status was -.26. These correlations indicated significantly different ($p < .05$) responses of minority and nonminority students.

The second listening study confirmed a positive correlation between listening ability and verbal ability and suggested that differing responses of minority and

nonminority students in the pilot study may be attributed to differing levels of verbal ability and not to problems of racial/ethnic bias in the items.

In summary, Spearritt's (1961) factor analysis lends support to the theory that listening comprehension is a separate factor that can be measured objectively. From an extensive review of studies investigating the correlations between listening comprehension and intelligence, Devine (1978) concluded that the observed relationships were small enough to indicate that listening depends upon factors other than intelligence. Mead's (1978) NAEP/SCA studies confirmed a high positive correlation between listening and verbal ability.

For the present study, this research is useful in that it supports the theory that listening comprehension, though related to intelligence, is a distinct factor that can be measured objectively.

Recommendations for the Development of Listening Comprehension Tests

Theorists have not agreed upon a conceptual framework for listening. This has led to considerable criticism of measurement attempts and has spurred recommendations for test construction.

The Massachusetts State Department of Education reviewed 37 listening instruments and found none that met their criteria. They listed three major problems with the

instruments they reviewed: concurrent validity, standardization, and inter-rater reliability (Brown et al., 1979).

After reviewing 107 articles on listening, 38 reports of listening research and ten listening tests, Faires (1980) concluded, "There is a great need for satisfactorily validated and reliable listening tests for all grades and age levels including adults" (p. 17). He added that there was not a single valid and reliable regionally or nationally normed listening test.

Two years later, Rubin et al., (1982) evaluated 45 published tests of speaking and listening for the Speech Communication Association (SCA). They concluded that these tests did not offer satisfactory solutions to measurement problems and noted that test/retest reliability was the most glaring inadequacy.

In discussing the development of their own test, Watson and Barker (1985) enumerated criticisms of existing tests. They included questionable validity, generalizability, norming procedures, passage length, methods of administration, mode of presentation, timing, test item difficulty, and sex and cultural bias.

Watson and Barker gave a brief history of the many recommendations for listening test construction at the SCA's annual meeting in 1985. In 1950, Caffrey identified three basic problems in evaluating listening--elimination of irrelevant factors such as attention, isolation of test

items from reading, and standardization of methods of presentation. In 1958, McCarthy emphasized the importance of controlling for content familiarity. He indicated that since listening is closely related to general information, individuals who are most familiar with the stimulus material will tend to have higher listening comprehension scores. In 1966, Langholz stressed that listening tests needed more discriminating items in order to ensure that they were reliable across grade levels. In 1980, Bostrom and Waldhart suggested that listening tests differentiate among short-term listening, short-term listening with rehearsal, and lecture listening.

Watson and Barker (1985) also reported on a project by Backlund et al. (1982) who were employed by the state of Massachusetts to conceive, implement, and evaluate basic skills in listening and then to establish criteria by which to judge published listening tests. Backlund, et al. selected four criteria:

1. Stimulus requires responder to perform as a listener.
2. Instrument distinguishes listening performance from reading and writing.
3. Instrument is free of gender, cultural, racial, ethnic content, and/or stereotyping.
4. Instrument assesses presence or absence of skills. (p. 2)

Of the seven listening tests these researchers evaluated, none met their predetermined cutoff point of 50% of the skills and criteria, so they provided recommendations for the development of future listening tests:

1. Use taped stimulus materials.
2. Use spoken, not written passages that are read aloud.
3. Require simple and minimal responses.
4. Produce the stimulus messages and questions on tape.
5. Provide test booklets with items in writing.
6. Use short stimulus materials (30 seconds to three minutes.
7. Use interesting stimulus materials (meaningful/ real-life).
8. Control vocalizations used in the stimulus materials. (p. 2)

Mead and Rubin (1985) provided these recommendations for listening comprehension tests:

1. The stimuli should represent typical oral language and not consist of passages designed to be written material.
2. The stimuli should model the language that students might hear in the classroom.
3. The passages should be interesting and short since listening performance is influenced by motivation and memory.

4. The passages should consist of topics grounded in experience common to all students.

5. Multiple-choice items should focus on the most important aspects of the passage rather than trivial details and be derived from the passage, without reliance on student's prior knowledge or experience. (p. 2)

They also suggested that examinees might be required to select a picture or perform a task as an alternative to multiple choice items. The authors omitted a possible third alternative of requiring students to give a verbal response.

Theorists have not agreed whether listening is a unitary skill or a series of subskills. Consequently, a critical question for the test developer is whether a test of listening comprehension should contain subtests or yield a single raw score.

Binford (1977) attempted to answer this question by administering an experimental battery of eight listening comprehension subtests to 396 students in fifth and sixth grades in Mason City, Iowa. He investigated the uniqueness of the subtests in the battery by examining the corrected correlations, reliabilities of difference, and factor analysis. These procedures indicated little evidence of uniqueness among the subtests and only one listening factor. While Binford did not make recommendations for test design, he stated, "The results of this study lend little or no support to providing separate subtest scores for different

facets or dimensions of listening comprehension, or for constructing a diagnostic battery of listening comprehension skills" (p.173-174).

In summary, several theorists have made recommendations for the development of listening tests in response to the need for a suitable instrument. They have primarily offered suggestions concerning the nature of the stimulus materials and, in some cases, have offered a rationale for their recommendations. Binford's (1977) factor analysis lent support to the construction of a listening comprehension test which yields a single raw score.

Selected recommendations were incorporated into the listening comprehension test developed in this study.

Instructional Implications

A number of researchers have investigated variables which seem to influence listening comprehension. Woodcock and Clark (1968) investigated the differences in listening comprehension among elementary school children (aged 9-4 to 11-3) in Nashville, Tennessee, who listened to recorded passages presented at different rates ranging from 78 words per minute to 428 words per minute and answered 28 multiple choice questions about the passages. One hundred sixty-two children from three levels of intelligence, as measured by the Peabody Picture Vocabulary Test, comprised the sample. Results indicated that listening rates of 228 to 328 words per minute were more efficient for learning and retention

than the normal rate of 178 words per minute. Subjects with lower IQ's performed better at rates which were slower than the most efficient rates for higher IQ subjects. According to Woodcock and Clark, their study provided evidence that high-speed listening can be an efficient learning medium for elementary school children.

Lundsteen (1971) noted that the use of compressed speech in instruction may be the future answer to problems of attention. She explained that the time difference between speed of speaking and speed of thought allows the listener's attention to wander. Therefore, if the speech rate were increased, the listener's mind would not wander from the message.

Schunk and Rice (1983) investigated the effects of self-verbalization of listening comprehension strategies on children's cognitive skills and self-efficacy. A total of 42 language-deficient second, third and fourth graders from seven schools within one school district participated. Pretest and Posttest instruments were the Science Research Associates (SRA) Achievement Series (listening comprehension portion) and an efficacy scale on which subjects predicted their success on the SRA questions. The children were randomly assigned within grade and school to one of two experimental conditions, strategy self-verbalization or no strategy verbalization. Training for the strategy verbalization group involved the teacher reading a story aloud to the group, verbalizing five strategies and

prompting the children to say each strategy aloud. The no-strategy group did not repeat strategies after the teacher verbalized them.

Results indicated that self-verbalization led to a higher level of self-efficacy across grades and enhanced older children's listening comprehension skills but did not promote listening comprehension among second graders. Schunk and Rice (1983) indicated that their study had teaching implications for providing strategy verbalization instruction to children whose past listening comprehension performances were deficit.

The construction view of comprehension states that comprehension of material is based on an active interplay between the person's cognitive structures (schemata) relevant to the material being comprehended and elements of the material itself (Townsend, 1983). According to this theory, an appropriate schema must first be activated to understand the incoming material. When the theme of the material changes, there must be a corresponding change in schemata or the material will not be understood.

With this view in mind, Townsend (1983) examined the ability of children to make appropriate schema shifts to accompany shifts in the subject matter of taped passages. The subjects were 40 third graders and 40 sixth graders, with a mean age of 8-7, in upstate New York. They listened to two taped passages and were asked to repeat what they had heard. Children in the cued shift condition were explicitly

cued to the shift in passages by twice repeating the title of the passage just prior to hearing. Children in the uncued shift condition heard both titles (twice) before hearing the first passage.

Townsend (1983) indicated that the findings of his study showed that children from both the third and sixth grades have difficulty with schema shifting in a listening comprehension task. Also, the children's recall of the second passage was markedly reduced when the shift in passages was not explicitly cued.

Just because researchers have demonstrated that a certain variable influences listening comprehension does not mean that teachers should start teaching it. Pearson and Fielding (1983) commented on this. They stated:

> It is one thing to be able to demonstrate that students with a better story schema understand stories better than those with a weaker story schema; it is quite another to demonstrate that providing those who are weak with a stronger story schema now comprehend better. (p. 18)

Pearson and Fielding (1983) summarized the research in instructional approaches for the teaching of listening comprehension and listed five conclusions based on their review:

1. Listening training in the same skills typically taught in reading comprehension curricula tends to improve listening comprehension.

2. Listening comprehension is enhanced by various kinds of active verbal responses on the part of students during and after listening.

3. Listening to literature tends to improve listening comprehension.

4. Certain types of instruction primarily directed toward other areas of the language arts may improve listening comprehension as well.

5. The direct teaching of listening strategies appears to help children to become more conscious of their listening habits than do more incidental approaches. (p. 9)

In summary, researchers have investigated variables which influence listening comprehension, such as compressed speech, self-verbalization strategy, and activation of schemata and have indicated that their findings have teaching implications. In their summary of the research, Pearson and Fielding (1983) tentatively concluded that listening comprehension can be improved through fairly direct instructional strategies.

Although, this literature does not relate directly to the present study, it points to the need for a valid and reliable test which can identify those children who will benefit most from listening comprehension instruction.

CHAPTER III

METHODOLOGY

This chapter contains a detailed description of the research plan used in this study. Included are descriptions of the samples, test construction methods, procedures, and statistical analyses.

Samples

The samples for this study included 570 children in four groups. Two field test samples contained 101 (45 boys and 56 girls across ages 5-12) and 106 (52 boys and 54 girls across ages 5-12) children. A non-referred (standardization) sample contained 308 (154 boys and 154 girls across ages 5-12) children. A referred sample contained 55 (32 boys and 23 girls across ages 6-12) children. All samples were randomly drawn from the population of school children in grades K-6 who attend schools in Edmond, Oklahoma, and the surrounding area (approximately 100 mile radius). This area included small rural, inner city, suburban and private schools, thereby representing a wide range of socio-economic levels.

Design

This study is categorized as test development research. It is not an experimental or quasi-experimental design. It is the development of an instrument that can be used in both an individual and group testing setting. The instrument was standardized and representative z-scores and percentiles were obtained.

The test has one form across grade levels K-6. It is administered verbally and requires a verbal response when administered individually or written response when administered to a group. The verbal and written response modes are ideal for assessing listening comprehension. The effect of an examinee's reading ability on listening performance is not an intervening variable. Also, these modes test recall, rather than recognition, which is consistent with the definition of listening comprehension.

The effect of an examinee's writing ability may be an intervening variable in the written response mode. This variable was controlled in the following ways:

1. The questions call for short and simple responses requiring a minimum of writing.

2. Answers containing spelling and letter formation errors were not scored as incorrect.

3. The written response mode was not used below third grade.

The effect of this variable was also investigated by statistically comparing the scores of group-tested and

individually-tested children. The results of this analysis is reported in chapter four.

The test contains five original passages with eight to ten questions for each, totaling 50 items. A 50 item test resulted in a relatively short administration time, thereby minimizing fatigue and allowing the test to be given with other tests. (Administering more than one test in a session is typical practice when a child is individually evaluated.) Including 50 items also maximized the variability on the test, thus improving discrimination among examinees' levels of performance.

Item Development

During item development, ten passages and 200 questions were written, with each passage serving as the stimulus for 20 questions (appendix A). These questions were divided into two sets of ten, were approximately parallel in difficulty, and ordered as the information appeared in the passage. This overproduction of passages and items was designed to facilitate selection of the best five passages and 100 questions.

Content validity was built into the test by using the following criteria when writing the passages and questions:

1. The passage material was randomly selected from a larger sample of material written according to the criteria stated below in items 2 through 5.

2. The passages contain only material the examinees would not be likely to know without having heard the passages. In this way, the examinees' knowledge of general information ought not to be a factor in performance.

3. The passages and questions contain typical oral language familiar to school children to minimize the effects of vocabulary knowledge.

4. The passage material was selected to minimize bias (age, gender, cultural, ethnic, and racial).

5. The questions call for simple and minimal responses, so that examinees' verbal or writing skills will affect performance minimally.

During the item development stage, 25 teachers were asked to evaluate the passages. The evaluator materials contained an evaluation form with a letter explaining the task (appendix B) and the passages and questions. Teachers were specifically asked to rank the passages and comment on them. They were also asked a question regarding face validity, "Does this test, overall, appear to measure listening comprehension?" The results of their responses are reported in chapter four.

Field Test

The five passages receiving the best rankings, along with their 100 questions, were selected for field testing. This initial pool of items was field tested in two forms, A and B, so that any one examinee was required to respond to

only half of the items. Each form contained the same five passages and 50 questions. Form A was administered to 101 children and form B was administered to 106 children (field test samples). The tests were individually administered (verbal response mode) to children in grades K-2 and group administered (written response mode) to children in grades 3-6. All tests were administered and scored by the author of the study.

Field test materials (appendix C) included a parent permission letter, directions for administering and scoring, the five selected passages, and three protocols: one for group administration, one for individual administration of form A, and one for individual administration of form B. The individual protocols listed the correct response, or the scoring criteria, in parentheses following each question. Items were scored "pass" or "fail" accordingly.

The field test served two purposes. First, it provided an opportunity to evaluate the test's administration and scoring procedures. The examinees' responses given during the field test were used to develop a scoring guide. Second, the field test provided an opportunity to select the best 50 items, through quantitative evaluation, for inclusion on the final form of the test.

Data Analyses

Factor analysis was conducted independently on the two forms of the field test data. Only items that loaded

saliently (.300 or higher) on the first factor from each analysis were considered for selection.

Item analysis was also conducted independently on the two forms. Because both the validity and reliability of a test depends greatly on the characteristics of its items, item analysis made it possible to ensure high test reliability and validity. Item analysis provided difficulty and discrimination values for the items. It was expected that only items with a moderate spread (.20 -.80) of difficulty would be selected and that, from those remaining, the 50 best discriminating items could be selected. However, this criteria reduced the number of items to below 50. Items with a discriminating value of .35 or higher were selected and the difficulty range was expanded to .32 to .96. The results of these analyses are reported in chapter four.

Standardization

The final form of the test was compiled by combining the same five passages used in the field study and the 50 selected items. Standardization materials (appendix D) included a new parent permission letter, revised directions for administering and scoring, and two new protocols for group and individual administration.

The test was administered to the non-referred sample. Children in grades K-2 were tested individually and children in grades 3-6 were tested individually and in groups. The

tests were administered and scored by the author of the study and by graduate students who are studying to be psychometrists and school psychologists. The graduate students were trained in administration and scoring procedures.

Data Analysis

The non-referred sample data were examined through item analysis. This allowed the test items to be characterized in terms of their difficulty levels and discrimination values.

Three kinds of reliability were investigated--internal consistency, test-retest, and inter-examiner. Concurrent, and factorial validity were also investigated. Standard scores were established using the scores from the non-referred sample. Internal consistency reliability was estimated by computing a Kuder-Richardson 20 coefficient. Test-retest reliability was investigated by administering the test to children across grade levels in the non-referred sample, allowing one week to elapse, administering the test a second time, and computing the Pearson correlation between test and retest scores. Some factors typically affecting reliability were controlled. These included making the item format consistent across grade levels, requiring recall of information, and providing standardized directions. Inter-examiner reliability was investigated by asking different examiners to score the same tests, which had been

administered and audiotaped by another examiner, and calculating the correlations between original and rescored responses. Scoring subjectivity was minimized by use of the scoring guide which had been developed from field test responses.

Concurrent validity was investigated by administering the test to groups of children in the fourth and sixth grades, and determining the correlation between the test scores and listening in a classroom situation. The criterion, classroom listening, contained the following components:

1. Each teacher selected criterion material and verbally presented this material to his/her class. The selected material was consistent with the curricula and new to the students.

2. Each teacher tested over the selected material directly following his/her presentation. The (teacher's) test required recall of information, and called for a written response. The (teacher's) test also included a question which asked if the student had heard the material before. The intent of this question was to provide a means by which the effects of previous exposure to the material could be statistically removed from the relationship between listening comprehension and classroom listening. This method of investigating concurrent validity validated the written response mode.

The fourth grade teacher presented a science lesson on earthquakes and volcanoes and the sixth grade teacher presented a history lesson on Greece. Each teacher asked eleven questions. The eleventh question asked if the student had heard the material before. These lessons and questions are presented in appendix E.

Factorial validity was investigated through factor analysis of the referred sample scores. The test could then be characterized in terms of its major factor. Since only test items that loaded on a single factor were included on the test, it was expected that only one factor would be salient.

To determine if there were any significant administration-type differences, t-tests were conducted between the scores of group-tested and individually-tested children in the third through sixth grades.

The non-referred sample data were analyzed by analysis of variance with multiple contrasts to determine where real differences existed, if any, in age, and gender groups. Groups that did not differ significantly were combined. The normative values for each group were determined and assigned a standard score (z-score) and percentile.

To predict how an examinee's score may vary on retesting, confidence limits were determined by computing the test's standard errors of measurement which could be used to establish confidence intervals. This allows an

examinee's score to be expressed as a range of scores rather than an exact point.

The final form of the test was individually administered to the referred sample. To determine whether there were any significant differences in listening comprehension between referred and non-referred students, t-tests were calculated on samples matched by age and gender. The results of all analyses are reported in chapter four.

CHAPTER IV

RESULTS

This chapter contains a description of the findings of this study. Included are the results of the item development, field test, and standardization procedures, the reliability and validity investigations, and comparison of referred and non-referred samples.

Item Development Results

During item development, ten passages and 200 questions were written. Twenty-five teachers were asked to evaluate these materials by ranking the passages from 1 to 10, with 1 being the best passage for measuring listening comprehension and 10 being the worst. They were also asked to respond to a question regarding face validity.

Table I shows the rankings and totals of the 17 teachers who responded. Five passages--alphabet, animals, folk tale, food, and plants--received the highest rankings. All 17 teachers gave a "yes" response to the question, "Does this test, overall, appear to measure listening comprehension?"

The selection of five passages reduced the number of associated questions to 100. These 100 questions were

TABLE I

RANKINGS OF PASSAGES BY 17 TEACHERS

Passage	Rank																	Total
ALPHABET	8	8	6	4	1	5	8	1	9	4	8	5	7	3	3	9	1	90
ANIMALS	4	7	2	9	9	3	5	6	3	5	3	6	4	5	5	6	3	85
COMPUTERS	10	3	7	3	4	2	4	10	10	10	5	4	5	9	4	5	5	100
FOLK TALE	6	2	5	1	8	4	6	2	2	1	1	10	6	4	1	3	6	68
FOOD	3	1	10	2	6	8	1	4	4	7	2	2	10	6	2	4	7	79
GLASS	1	4	9	6	3	10	9	5	6	3	9	8	2	10	8	1	9	103
PLANTS	9	10	1	8	2	1	10	3	1	2	4	1	1	1	10	10	10	84
SHELLS	2	9	4	10	10	9	3	9	8	8	7	9	9	8	7	2	4	118
STARS	5	6	3	5	5	6	2	8	7	6	6	7	8	2	6	7	8	97
THE WHEEL	7	5	8	7	7	7	7	7	5	9	10	3	3	7	9	8	2	111

divided into two sets of 50. Two forms, A and B, were constructed. Each form contained the same five passages and 50 questions (items).

Field Test Results

Each form of the test was administered to one of the field test samples. The tests were individually administered to children in grades K-2 and group administered to children in grades 3-6.

Descriptive statistics were calculated for the field test data. Tables II and III contain the raw score means, standard deviations and sample sizes for forms A and B, respectively.

Each set of field test data was factor analyzed by principal component analysis. Because the first principal component is the linear combination of items that account for the largest amount of variance in the sample, items with component loadings of .30 or higher on the first component were selected for possible inclusion on the final form of the test. Difficulty and discrimination values were also calculated for each set of field test data. From this group, the 50 best items discriminating items (.35 or higher) within the desired range of difficulty (.32 to .96) were selected. Tables IV and V show the results of the factor and item analyses for the selected items.

A scoring guide was written using field test responses. Examples of "pass" and "fail" responses and

TABLE II

DESCRIPTIVE RAW SCORE STATISTICS FOR FORM A

Grade		Boys	Girls	Total
K	Mean	31.20	26.50	28.62
	SD	8.50	3.45	6.39
	n	5	6	11
1	Mean	32.80	37.80	35.30
	SD	11.90	6.06	9.29
	n	5	5	10
2	Mean	35.20	36.20	35.70
	SD	3.56	3.90	3.56
	n	5	5	10
3	Mean	35.29	32.10	33.41
	SD	5.77	5.97	5.93
	n	7	10	17
4	Mean	39.60	31.14	36.12
	SD	3.86	4.53	5.87
	n	10	7	17
5	Mean	40.82	39.09	39.96
	SD	6.56	3.59	5.23
	n	11	11	22
6	Mean	41.00	42.67	42.43
	SD	2.83	4.64	4.38
	n	2	12	14
Total n		45	56	101

TABLE III

DESCRIPTIVE RAW SCORE STATISTICS FOR FORM B

Grade		Boys	Girls	Total
K	Mean	30.40	24.60	27.50
	SD	5.41	8.62	7.44
	n	5	5	10
1	Mean	30.20	25.40	27.80
	SD	15.87	7.57	11.99
	n	5	5	10
2	Mean	30.50	40.75	36.63
	SD	8.23	3.10	7.25
	n	4	4	8
3	Mean	40.33	33.60	36.79
	SD	6.06	6.85	7.19
	n	9	10	19
4	Mean	36.71	36.38	36.53
	SD	4.68	5.88	5.17
	n	7	8	15
5	Mean	40.36	41.00	40.63
	SD	6.07	5.83	5.85
	n	14	10	24
6	Mean	41.38	40.58	40.90
	SD	4.17	6.99	5.90
	n	8	12	20
Total n		52	54	106

TABLE IV

CHARACTERISTICS OF ITEMS SELECTED FROM FORM A

Item	Discrimination	Difficulty	Factor Loading
Alphabet			
3	.53	.72	.58
5	.59	.41	.59
Animals			
3	.41	.70	.38
4	.55	.64	.58
5	.35	.72	.32
10	.39	.93	.44
Folk Tale			
3	.59	.70	.63
5	.39	.54	.37
9	.39	.91	.42
Food			
2	.47	.81	.54
4	.64	.76	.70
5	.51	.79	.58
7	.41	.76	.45
8	.44	.72	.48
10	.46	.57	.50
Plants			
2	.47	.50	.48
6	.42	.66	.40
7	.46	.61	.45
8	.40	.68	.41
10	.50	.55	.44

TABLE V

CHARACTERISTICS OF ITEMS SELECTED FROM FORM B

Item	Discrimination	Difficulty	Factor Loading
Alphabet			
1	.49	.76	.49
2	.52	.37	.49
4	.36	.32	.31
5	.37	.54	.36
6	.35	.59	.32
7	.62	.61	.62
10	.37	.68	.36
Animals			
1	.43	.38	.42
2	.53	.59	.52
3	.45	.87	.46
5	.52	.34	.49
9	.44	.90	.46
Folk Tale			
1	.37	.43	.33
2	.45	.65	.43
3	.59	.52	.60
8	.44	.91	.47
10	.48	.96	.53
Food			
1	.44	.70	.48
4	.67	.72	.69
6	.56	.67	.55
8	.54	.86	.57
9	.54	.72	.54
10	.54	.85	.56

TABLE V (Continued)

Item	Discrimination	Difficulty	Factor Loading
Plants			
2	.59	.83	.63
3	.57	.65	.57
4	.54	.67	.55
7	.42	.61	.40
8	.46	.87	.49
9	.64	.87	.68
10	.67	.76	.68

specific criteria for scoring the items were included. It was not necessary to revise administration procedures.

Standardization Results

The 50 items selected for inclusion on the test were combined in one form and administered to the non-referred sample. The tests were individually administered to children in grades K-2 and individually and group administered to children in grades 3-6.

Descriptive statistics were calculated for the non-referred sample data. Table VI shows the raw score means and standard deviations of total scores on the test. The data are reported for boys and girls, separately and combined, at each age level. Table VII shows both total and subtest (passage) means and standard deviations.

T-tests for independent samples were calculated to test for differences between administration-types. The results indicated only non-significant administration-type differences. Administration types were combined in subsequent analyses.

T-tests for independent samples were also calculated to test for gender differences within each age level. Results showed non-significant gender differences. Boys and girls scores were therefore combined in subsequent analyses.

A one-way analysis of variance was performed to test for differences between age groups. A significant F statistic, ($F = 16.338$, $df = 7, 300$, $p < .001$), was

TABLE VI

DESCRIPTIVE RAW SCORE STATISTICS

Age		Boys	Girls	Total
5	Mean	20.00	14.00	16.00
	SD	3.61	3.52	4.47
	n	3	6	9
6	Mean	25.00	18.57	21.54
	SD	7.04	8.54	8.26
	n	6	7	13
7	Mean	26.00	30.40	28.00
	SD	8.15	2.61	6.42
	n	6	5	11
8	Mean	33.29	33.28	33.28
	SD	8.10	9.47	8.90
	n	17	29	46
9	Mean	34.52	32.37	33.42
	SD	7.65	7.86	7.77
	n	29	30	59
10	Mean	33.52	32.12	32.86
	SD	7.35	7.41	7.34
	n	29	26	55
11	Mean	37.34	37.47	37.40
	SD	8.22	7.31	7.75
	n	38	34	72
12	Mean	37.96	36.71	37.47
	SD	6.04	9.40	7.47
	n	26	17	43
Total n		154	154	308

TABLE VII

DESCRIPTIVE RAW SCORE STATISTICS FOR PASSAGES AND TOTAL TEST (n = 308)

Passage	Mean	Standard Deviation
Alphabet	6.35	2.01
Animals	5.09	2.28
Folk Tale	5.52	1.91
Food	8.88	2.32
Plants	7.75	2.62
Total	33.59	8.98

obtained. Tukey's HSD multiple comparisons were calculated to pinpoint the differences. Table VIII shows the results of these analyses.

Box plot displays (Figure 1) were generated for total score at each age group across ages 5-12. Box plots are graphical analogs to the one-way analysis of variance, although they use rank order statistics instead of means. The minimum and maximum score values of 7 and 50 across all ages form the basis of the display. Medians are marked with a plus (+) sign. The lower and upper hinges are at the left and right edges of the boxes. The ends of the horizontal lines denote the outermost values. Outliers are marked with an asterisk (*).

Z-scores and percentile ranks were calculated for each age level. They are shown in table IX.

The non-referred sample data were factor analyzed and discrimination and difficulty values were calculated for each of the 50 items. Table X shows the results of these analyses.

Reliability

Three kinds of reliability coefficients were calculated--internal consistency, test-retest, and inter-examiner. Internal consistency reliability was estimated by calculating a Kuder-Richardson 20 coefficient for each age level. Table XI shows these coefficients and the associated standard errors of measurement for each age.

TABLE VIII

SUMMARY STATISTICS FOR AGE GROUP DIFFERENCES

Analysis of Variance

Source	SS	Df	Mean Square	F	Prob.
Between	7589.84	7	1084.26	16.338	.000
Within	19909.16	300	66.36		

Tukey HSD Multiple Comparisons

Matrix of pairwise absolute mean differences between ages

Age	5	6	7	8	9	10	11	12
5	0							
6	5.78	0						
7	*12.87	7.09	0					
8	*18.26	*12.48	5.39	0				
9	*18.49	*12.71	5.62	.23	0			
10	*17.85	*12.07	4.98	.41	.64	0		
11	*22.72	*16.94	*9.85	4.47	4.23	4.87	0	
12	*22.68	*16.91	*9.82	*4.43	4.20	*4.83	.04	0

* $p < .05$

FIGURE 1

BOX PLOTS OF TOTAL SCORES ACROSS AGES 5-12

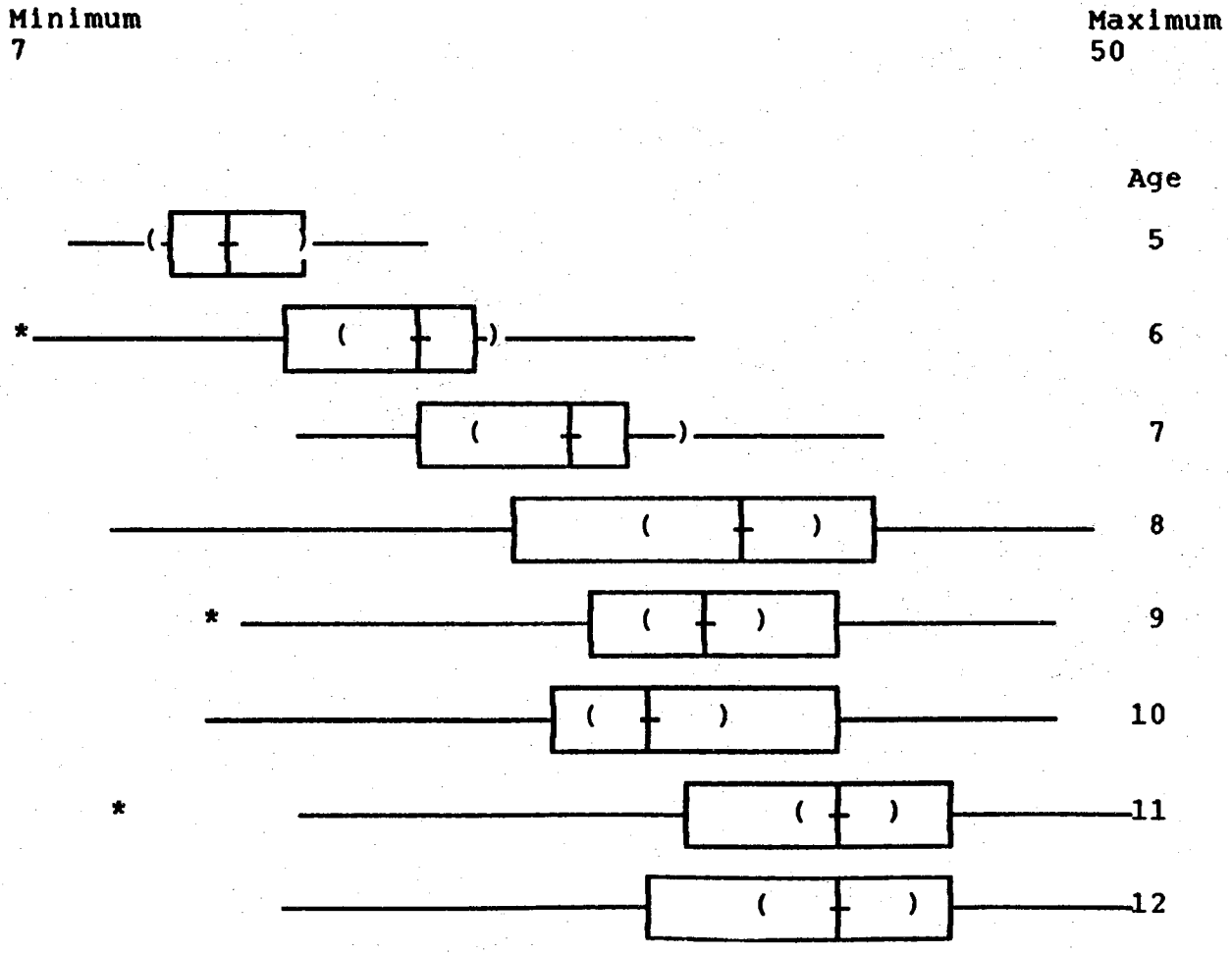

TABLE IX

Z-SCORES AND PERCENTILES FOR AGES 5-12

Raw Score	Z-score	Percentile
Age 5 (n = 9)		
10	-1.342	22.22
14	-0.447	33.33
15	-0.224	44.44
16	0.000	66.67
19	0.671	77.78
21	1.118	88.89
23	1.565	99.00
Age 6 (n = 13)		
7	-1.760	7.69
8	-1.639	15.38
15	-0.791	23.08
18	-0.428	30.77
19	-0.307	38.46
21	-0.065	46.15
23	0.177	61.54
25	0.419	76.92
30	1.024	84.62
33	1.387	99.00
Age 7 (n = 11)		
19	-1.402	9.09
20	-1.246	18.18
21	-1.091	27.27
25	-0.467	36.36
27	-0.156	45.45
29	0.156	54.55
31	0.467	81.82
34	0.935	90.91
40	1.870	99.00

TABLE IX (Continued)

Raw Score	Z-score	Percentile
Age 8 (n = 46)		
12	-2.392	2.17
17	-1.830	4.35
18	-1.718	6.52
19	-1.605	10.87
20	-1.493	13.04
23	-1.156	15.22
24	-1.043	17.39
25	-0.931	21.74
26	-0.819	23.91
27	-0.706	28.26
29	-0.481	30.43
30	-0.369	32.61
31	-0.257	43.48
33	-0.032	47.83
35	0.193	52.17
36	0.305	56.52
37	0.418	60.87
38	0.530	65.22
39	0.643	69.57
40	0.755	76.09
41	0.868	80.43
42	0.980	86.96
43	1.092	91.30
45	1.317	95.65
47	1.542	97.83
48	1.654	99.00
Age 9 (n = 59)		
15	-2.372	1.69
17	-2.114	5.08
20	-1.728	6.78
21	-1.599	10.17
22	-1.470	11.86
23	-1.342	15.25
24	-1.213	16.95
26	-0.956	18.64
27	-0.827	20.34
28	-0.698	22.03
29	-0.569	25.42

TABLE IX (Continued)

Raw Score	Z-score	Percentile
Age 9 (n = 59)		
30	-0.441	33.90
31	-0.312	35.59
32	-0.183	38.98
33	-0.055	49.15
34	0.074	50.85
35	0.203	52.54
36	0.332	61.02
37	0.460	66.10
38	0.589	71.19
39	0.718	79.66
40	0.847	81.36
42	1.104	84.75
43	1.233	93.22
44	1.361	98.31
47	1.748	99.00
Age 10 (n = 55)		
15	-2.432	1.82
19	-1.887	3.64
20	-1.751	5.45
23	-1.342	9.09
24	-1.206	10.91
25	-1.070	18.18
26	-0.934	20.00
27	-0.797	25.45
29	-0.525	36.36
30	-0.389	41.82
31	-0.253	47.27
32	-0.116	52.73
33	0.020	54.55
35	0.292	56.36
36	0.428	67.27
37	0.565	72.73
38	0.701	74.55
40	0.973	85.45
41	1.109	89.09
42	1.246	90.91
43	1.382	92.73
44	1.518	94.55

TABLE IX (Continued)

Raw Score	Z-score	Percentile
Age 10 (n = 55)		
45	1.654	96.36
47	1.927	99.00
Age 11 (n = 72)		
12	-3.278	1.39
19	-2.375	2.78
20	-2.246	4.17
21	-2.117	6.94
25	-1.601	8.33
27	-1.342	13.89
31	-0.826	18.06
32	-0.697	22.22
33	-0.568	26.39
34	-0.439	27.78
35	-0.310	31.94
36	-0.181	38.89
37	-0.052	43.06
38	0.077	45.83
39	0.206	54.17
40	0.335	59.72
41	0.464	63.89
42	0.593	69.44
43	0.722	86.11
44	0.851	87.50
45	0.980	90.28
46	1.109	91.67
47	1.239	93.06
48	1.368	95.83
49	1.497	97.22
50	1.626	99.00

TABLE IX (Continued)

Age 12 (n = 43)		
Raw Score	Z-score	Percentile
18	-2.606	2.33
22	-2.071	4.65
25	-1.669	6.98
26	-1.535	11.63
30	-1.000	16.28
31	-0.866	23.26
32	-0.732	27.91
33	-0.598	32.56
35	-0.330	34.88
36	-0.196	37.21
37	-0.062	44.19
38	0.072	48.84
39	0.206	55.81
40	0.339.	62.79
41	0.473	65.12
42	0.607	69.77
43	0.741	74.42
44	0.875.	81.40
45	1.009	88.37
47	1.277	97.67
50	1.678	99.00

TABLE X

CHARACTERISTICS OF ITEMS

Item	Discrimination	Difficulty	Factor Loading
Alphabet			
1	.29	.83	.35
2	.47	.40	.46
3	.48	.73	.49
4	.47	.34	.47
5	.48	.47	.49
6	.35	.51	.35
7	.46	.56	.44
8	.47	.63	.49
9	.40	.62	.42
Animals			
1	.49	.58	.49
2	.50	.69	.52
3	.51	.68	.51
4	.45	.74	.47
5	.32	.87	.38
6	.38	.75	.41
7	.41	.47	.39
8	.31	.85	.31
9	.12	.71	.09
Folk Tale			
1	.40	.53	.38
2	.48	.71	.51
3	.48	.80	.49
4	.50	.52	.48
5	.42	.55	.42
6	.44	.72	.45
7	.42	.75	.42
8	.31	.95	.37
Food			
1	.44	.70	.48
2	.31	.91	.34
3	.46	.84	.49
4	.49	.77	.52
5	.33	.84	.36
6	.25	.53	.25

TABLE X (Continued)

Item	Discrimination	Difficulty	Factor Loading
Food			
7	.32	.68	.29
8	.46	.71	.48
9	.31	.77	.33
10	.46	.49	.46
11	.42	.86	.46
12	.41	.84	.44
Plants			
1	.27	.95	.35
2	.33	.33	.35
3	.50	.54	.48
4	.41	.65	.39
5	.50	.65	.51
6	.31	.46	.27
7	.29	.43	.23
8	.48	.74	.49
9	.32	.91	.34
10	.44	.86	.48
11	.49	.46	.48
12	.57	.78	.58

TABLE XI

KUDER-RICHARDSON FORMULA 20 COEFFICIENTS OF INTERNAL CONSISTENCY AND STANDARD ERRORS OF MEASUREMENT

Age	Reliability	SEM	n
5	.63	2.72	9
6	.89	2.74	13
7	.79	2.94	11
8	.90	2.81	46
9	.87	2.80	59
10	.85	2.84	55
11	.89	2.57	72
12	.88	2.59	43
Total			308

Test-retest reliability data were obtained by retesting 169 children after a one week interval and correlating their scores. Table XII shows these test-retest coefficients for each age level. Paired t-tests were also performed on the test and re-test means to investigate test differences due to practice effects. Significant differences (gains) were observed for ages 5-10. The results of the paired t-tests are shown in table XIII.

Inter-examiner reliability coefficients were obtained by correlating the scores assigned by different examiners to the same tests. A total of 170 tests were scored by two different examiners, using audiotapes of the test administrations. The results are displayed in table XIV.

Validity

Concurrent validity was estimated by calculating a Pearson correlation between the classroom listening scores and total scores of the listening comprehension test. Because only two children indicated they had heard the classroom listening materials before testing, it was not possible to covary prior learning. Seventeen fourth graders and 28 sixth graders participated in the classroom listening testing. Pearson correlations were significant ($r = .755$, $df = 15$, $p < .01$) for fourth grade and ($r = .391$, $df = 26$, $p < .05$) for sixth grade.

TABLE XII

TEST-RETEST RELIABILITIES

Age	Reliability	n
5	.781	7
6	.981	10
7	.785	7
8	.883	36
9	.874	36
10	.796	29
11	.868	31
12	.734	13
Total		169

TABLE XIII

PAIRED T-TESTS ON TEST AND RETEST MEANS

Age	n	t	df	Probability
5	7	-3.655	6	*.011
6	10	-5.706	9	**.000
7	7	-8.555	6	**.000
8	36	-7.128	35	**.000
9	36	-7.677	35	**.000
10	29	-2.691	28	*.012
11	31	-1.600	30	.120
12	13	-2.031	12	.065
Total	169			

* $p < .05$
** $p < .001$

TABLE XIV

INTER-EXAMINER RELIABILITIES

Age	Reliability	n
5	.985	8
6	.990	8
7	.983	9
8	.994	33
9	.989	32
10	.996	27
11	.977	38
12	.997	15
Total	.991	170

Referred Sample Results

The test was individually administered to 55 children across ages 6-12 in the referred sample. To determine whether there were any significant differences between referred and non-referred children on listening comprehension, a matched samples t-test was calculated, using a sample of non-referred children matched on age and gender. The group means were 32.52 (non-referred) and 25.22 (referred). The standard deviations were 8.47 (non-referred) and 11.30 (referred). A significant t statistic ($t = 3.53$, $df = 108$, $p < .001$) was obtained.

CHAPTER V

DISCUSSION

This chapter presents the summary, conclusions, and recommendations of this study.

Summary

The main purpose of this study was to develop a listening comprehension test, standardize the instrument and determine its reliability and validity.

A literature search was conducted which included three major topics--defining listening, the relationship between listening comprehension and other factors, and recommendations for development of listening comprehension tests. A brief review of major findings in instructional implications for the teaching of listening comprehension was also included.

The participants in this study were 570 children in grades K-6 who attend public schools in Edmond, Oklahoma, and the surrounding area. They were divided into four samples: two field test samples (n's = 101 and 106), a non-referred sample (n = 308), and a referred sample (n = 55).

So that content validity might be built into the test, five criteria were established. Briefly, these criteria required that the passage material be randomly selected from a larger sample of material, contain only material that the examinees would not be likely to know, contain typical oral language familiar to school children, and be selected to minimize bias. The last criterion required that the questions call for simple and minimal responses. Ten original passages and 200 questions were developed using these criteria. Teachers were asked to rank these passages, using the criteria. Based on the rankings, five passages (alphabet, animals, folk tale, food, and plants) and their associated 100 items were selected and used in field tests. The 100 items were divided into two sets of 50 and put into separate forms, A and B. Each form contained the same five passages. Children in grades K-2 were tested individually (verbal response mode) and children in grades 3-6 were tested in groups (written response mode).

Raw score means and standard deviations were calculated for boys and girls, both separately and combined, at each grade level, for each form of the test.

Each set of field test data was factor analyzed by principal component analysis. The factor loadings for selected items ranged from .32 to .70 on form A and from .31 to .69 on form B. Discrimination and difficulty values were also calculated for each set of field test data. Discrimination values for selected items ranged from .35 to

.64 on form A and from .35 to .67 on form B. Difficulty values on form A ranged from .41 to .93 and from .32 to .96 on form B.

Field test results were used to develop a scoring guide for the final form of the test. Subjectivity in scoring was minimized by listing "pass"/"fail" responses and providing specific guidelines for scoring.

The final form of the test was administered to the non-referred sample. Children in K-6 were tested individually and children in 3-6 were tested in groups.

Raw score means and standard deviations were calculated for boys and girls, both separately and combined, for each age level and for each passage and total test.

T-tests were used to test for differences between group and individual administration types. There were no significant differences. T-tests were also used to test for gender differences at each age level. Again, no significant differences were observed. These groups were subsequently combined for further analyses.

One-way analysis of variance was performed to test for overall differences between age groups. Results showed that age was a significant variable. Tukey's HSD tests were used to pinpoint specific, significant differences. These were found between age 5 and ages 7-12, between age 6 and ages 8-12, and between age 7 and ages 11-12. Age 8 was significantly different from ages 5,6, and 12; age 10 was significantly different from age 12.

Box plot displays were generated for total score across ages 5-12. These plots graphically showed that ages 5, 6, and 7 were different from each other while ages 8-9 and 10-12 tended to group together.

Standard scores were calculated for each age group. Ages 5, 6, and 7 had inadequate sample sizes for norming purposes. Representative z-scores and percentiles were obtained for children from 8 to 12 years of age.

The results of factor analysis of the non-referred data indicated that the test is unidimensional. A single factor accounted for 18.2 percent of the total explained variance. All but five items (animals 9, food 6 and 7, and plants 6 and 7) exhibited salient factor loadings of .30 or higher on the first factor.

Item analysis based on the non-referred sample's results yielded acceptable discrimination and difficulty values for most items. Only five items (alphabet 1, animals 9, food 6, and plants 1 and 7) yielded discrimination values below .30. Three of these same five items (animals 9, food 6, and plants 7) also yielded factor loadings below .30. A subjective evaluation of these items suggests that their poor performance may have been a result of the scoring criteria that were used.

An overall appraisal of the item characteristics (discrimination, difficulty, and factor loadings) clearly shows that the test items contribute to the measurement of listening comprehension.

Internal consistency, test-retest, and inter-examiner reliabilities were calculated at each age level. The Kuder Richardson 20 (K-R 20) formula yielded high coefficients (.79 to .90) for all ages except 5 (.63). The K-R 20 coefficients were used to calculate standard errors of measurement (SEM) for all age groups. SEMs were consistent, ranging from 2.57 to 2.94.

Test-retest reliabilities (.73 to .98) were both acceptable and comparable with other standardized tests. Paired t-tests were performed using test and retest means at each age level. The results indicated practice effects, with ages 5-10 performing significantly better on the retest.

Inter-examiner reliability coefficients were high, with the lowest being .97. This may be attributed to the specificity of the scoring guide and the careful use of this guide by the examiners.

Concurrent validity coefficients were .76 and .39 for fourth and sixth grades, respectively. These results were significant and encouraging, but they are probably unstable due to the low range of scores on the ten-item (teacher's) test. The difference between the two correlations may be attributed to the content differences in the classroom listening material. The only consistencies were that the materials were derived from the curricula, and were presented without visual cues.

A secondary purpose of this study was to determine whether there were any significant differences between populations of referred and non-referred students on listening comprehension. All children in the referred sample were individually tested. After testing, they were matched (by age and gender) with a same-size sample of children in the non-referred group. A matched samples t-test showed significant differences ($p < .001$). It appears unlikely that the population means of the two groups are equal.

Conclusions

The procedures used in the development of this test were consistent with good test development practices. The writing of the stimulus materials was guided by pre-established content validity criteria. A field study facilitated the refinement of the test by providing a means to evaluate performance of the items through factor and item analysis. A standardized method of presentation, including directions for administering, timing requirements, and scoring criteria, were developed. The test provides a direct means of assessing a child's relative level of functioning in listening comprehension. Standard scores are provided for interpretation.

The test is a single factor test which may be administered individually to children, ages 5-12, or to groups of children in grades 3-6. Support for a single

factor test was found in the research of Binford (1977) who extracted only one listening factor from a battery of eight listening comprehension subtests and Spearritt (1961) who found a listening comprehension factor confined within nine tests of listening comprehension.

The test developed in this study differs from reading tests by measuring listening without any reading on the part of the examinee. Several theorists have addressed the necessity of separating listening and reading. Schell (1981) found that listening and reading are not equal in achievement and Brown et al. (1979) cautioned that assessment instruments for listening should be different from those which measure reading. Faires (1980) emphasized that requiring reading on a listening test contaminates results.

The test has demonstrated reliability and validity as a test of listening comprehension. Criticisms of many existing listening tests have been in the area of inter-rater reliability (Brown et al., 1979) and test-retest reliability (Rubin et al., 1982).

This test was validated using classroom listening as the criterion. It did not attempt to present the typical language of the "real world" or to measure a child's listening performance in that "real world." Listening skills required inside and outside the classroom may be quite different. When classroom listening is measured it may not generalize to other listening opportunities. It may

not represent the level of listening skill a child exhibits when the place, speaker, subject matter and affect changes. Hymes (1964) alluded to this when he stated that "the listener has to consider who speaks to whom, why, where, and when if the interpretation of the message is to be adequate" (p. 8). Also, Lundsteen (1971) indicated that "emotion" effects listening and Rixon (1981) indicated that the listener must find a reason for listening. More research is needed to determine the relationship between listening and these variables. We may find that the listening performance of a child improves naturally when the subject matter is more interesting and relevant.

In this study, significant differences were observed between referred and non-referred children's performances on listening comprehension. Referred children, designated learning disabled, scored significantly lower on listening comprehension. One of the most important uses of this test may be to diagnose learning difficulties related to listening comprehension.

Recommendations

There are several steps which need to be covered before the test is sufficiently refined for general use as a standardized test of listening comprehension.

The scoring criteria for the five test items which did not perform well (factor loadings below .300) need to be rewritten, or new items must be found. The protocols which

were administered to the non-referred group may, with new scoring criteria, be rescored using the new guidelines. Additional children at the lower ages, 5, 6, and 7 should be tested to expand the norming base.

Additional concurrent validity data should be collected across grades 3 through 6. A method of collecting validity data for young children, who are in grades K-2, should be devised and used to validate the test for the verbal response mode.

Using the other five passages which were developed at the beginning of the study, the procedures used in this study should be repeated to develop a second parallel form of the test.

REFERENCES

Backlund, P. M., Brown, K. L., Gurry, J., & Jandt, F. (1982). Recommendations for assessing speaking and listening skills. Communication Education, 31, 9-17.

Binford, F. (1977). A study of interrelationships among different approaches to measuring listening comprehension. Dissertation Abstracts International, 38, 07A.

Bostrom, R. N. & Waldhart, E. S. (1980). Components in listening behavior: The role of short-term memory. Human Communication Research, 6, 222-27.

Brown, C. (1965). Three studies in the listening of children. Speech Monographs, 32, 129-138.

Brown, D. (1954). Auding as a primary language ability. Unpublished doctoral dissertation, Stanford University.

Brown, K. L., Backlund, P., Gurry, J., & Jandt, F. (1979). Assessment of basic speaking and listening skills: State of the art and recommendations for instrument development. Boston: Massachusetts State Dept. of Education, Bureau of Research and Assessment. (ERIC Document Reproduction Service No. ED 178 970)

Caffrey, J. (1950). The establishment of auding-age norms. Education Digest, 15, 18-20.

Clark, H. & Clark E. (1977). Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovish, Inc.

Devine, T. (1978). Listening: What do we know after fifty years of research and theorizing? Journal of Reading, 296-304.

Devine, T. (1981). Listening skills assessment: Manual and script. Concord: New Hampshire State Dept. of Education, New Hampshire Educational Assessment Program. (ERIC Document Reproduction Service No. ED 236 651)

Duker, S. (1971). Listening: Readings II. Metuchen, NJ: Scarecrow Press.

Dunkel, P. (1986). Developing listening fluency in L2: Theoretical principles and pedagogical considerations. The Modern Language Journal, 70, 99-106.

Education Amendments of 1978. P. L. 95-561. 92 STAT 2201-2205.

Education For All Handicapped Children Act of 1975. P. L. 94-142. 34 C.F.R. 300

Faires, C. L. (1980, November). The development of listening tests. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 220 528)

Hymes, D. (1964). Toward ethnographics of communication: The analysis of communicative events. In P. P. Giglioli (ed.) Language and social context. London: Cox & Wyman Ltd.

Landolfi, L. (1984, July). Listening comprehension: Process, problems, and solutions. Paper presented at the Summer Meeting of Teachers of English to Speakers of Other Languages, Corvallis, OR. (ERIC Document Reproduction Service No. ED 253 108)

Langholz, A. P. (1966). A study of the relationship of listening test scores to test item difficulty. Dissertation Abstracts, 26, 6912.

Lundsteen, S. (1971). Listening: Its impact at all levels on reading and the other language arts. Urbana, IL: National Council of Teachers of English.

McCarthy, E. L. (1958). Auding ability related to achievement in two telecourses. Dissertation Abstracts, 18, 531-532.

Mead, N. (1978, November). Assessing listening ability: Relationships with verbal ability and racial/ethnic bias. Paper presented at the Annual Convention of the Speech Communication Association, Minneapolis, MN. (ERIC Document Reproduction Service No. ED 193 678)

Mead N. & Rubin D. (1985). Assessing listening and speaking skills. Washington, D. C.: National Inst. of Education. (ERIC Document Reproduction Service No. ED 263 626)

Pearson, P. D. & Fielding, L. (1983). Instructional implications of listening comprehension research. (Reading Education Report No. 39). Urbana: Illinois University, Center for the Study of Reading. (ERIC Document Reproduction Service No. ED 227 464)

Rixon, S. (1981). The design of materials to foster particular listening skills. The Teaching of Listening Comprehension, ELT Documents, 121 68-106.

Roberts, C. (1985, November). A user's response to the use of listening assessment instruments. Paper presented at the Annual Meeting of the Speech Communication Association, Denver, CO. (ERIC Document Reproduction Service No. ED 263 641)

Rubin, D. L., Daly, J., McCroskey, J., & Mead, N. (1982). A review and critique of procedures for assessing speaking and listening skills among preschool through grade twelve students. Communication Education, 31, 285-303.

Schell, L. (1981). The validity of the potential level via listening comprehension: A cautionary note. Paper presented at the Annual Meeting of the Plains Regional Conference of the International Reading Association, Des Moines, IA. (ERIC Document Reproduction Service No. ED ED 210 626)

Schunk, D. & Rice, J. (1983, August). Strategy verbalization effects on self-efficacy and listening comprehension. Paper presented at the Annual Meeting of the American Psychological Association, Anaheim, CA. (ERIC Document Reproduction Service No. ED 233 816)

Spearritt, D. (1961). A factorial analysis of listening comprehension. (Doctoral dissertation, Harvard University). American Doctoral Dissertations, 1961, 58.

Sticht, T. G., Beck, L., Hauke, R., Kleiman, G. & James, J. (1974). Auding and reading: A developmental model. Arlington, VA: Human Resource Research Organization.

Swanson, C. (1984, March). Monitoring student listening techniques: An approach to teaching the foundations of a skill. Paper presented at the Annual Meeting of the Eastern Communication Association, Philadelphia, PA. (ERIC Document Reproduction Service No. ED 240 653)

Townsend, M. (1983). Schema shifting: Children's cognitive monitoring of the prose-schema interaction in comprehension. Journal of Experimental Child Psychology, 36, 139-149.

Watson, K. & Barker, L. (1985, November). Listening assessment: The Watson-Barker Listening Test. Paper presented at the Annual Meeting of the Speech Communication Association, Denver, CO. (Eric Document Reproduction Service No. ED 264 613)

Woodcock, R. & Clark, C. (1968). Comprehension of a narrative passage by elementary school children as a function of listening rate, retention period, and IQ. The Journal of Communication, 18, 259-271.

APPENDIX A

ORIGINAL TEN PASSAGES WITH QUESTIONS

ORIGINAL TEN PASSAGES WITH QUESTIONS

ALPHABET

A long time ago people did not have an alphabet. They just drew little pictures when they wanted to write something. Each picture stood for a word. This did not work very well because they needed many pictures just to write a sentence.

The first people to use an alphabet were the Semites. The Semites lived about 3,500 years ago. They decided to write their language using signs. Their signs were like simple little pictures with one big difference. The signs did not stand for words. Each sign stood for a sound. For example, a picture of an ox head stood for the sound of "a." A house stood for the sound of "b." Waves of water was the sign for the "m" sound. The Semites discovered that they only needed about 30 signs or letters to write their language.

Other groups of people began to use the Semite alphabet. The Greeks and Romans changed the shapes of some of the letters. They gave up a few of the letters they did not need. They also added one or two letters of their own.

Today most languages are written with alphabets made from the one invented by the Semites long ago. A few languages are still written with signs that are not letters of an alphabet. Chinese writing, for instance, is done with such signs. But writing with an alphabet is much more common.

FORM A

1. Have people always had an alphabet?
2. In writing with pictures, what does each picture stand for?
3. Who were the first people to use an alphabet?
4. How long ago did the Semites live?
5. How did the Semites write the sound of "a"?
6. Waves of water was the sign for what sound?
7. Was the Semite alphabet ever changed?
8. Who changed the shape of the letters?
9. Today, do people use alphabets made from the one invented by the Semites?
10. Does Chinese writing use an alphabet?

FORM B

1. A long time ago, what did people do when they wanted to write something?
2. Writing with pictures does not work very well. Why?

3. What did the Semites invent?
4. In the Semite alphabet, what did the signs stand for?
5. How did the Semites write the sound of "b"?
6. How many letters were in the Semite alphabet?
7. What did the Greeks and Romans do to the Semite alphabet?
8. Were any letters ever added to the alphabet?
9. Are any languages written with signs that are not letters of an alphabet?
10. What is the most common way to write a language?

ANIMALS

Did you know that the world has so many animals that no one knows how many kinds there are? Scientists have found almost a million kinds of animals so far. Every year, hundreds of new kinds of animals are discovered.

Scientists have studied how animals are alike and how they are different. When they find that animals are alike, they put them into a group. They have divided all the animals into two main groups--animals without backbones and animals with backbones. Worms, starfish, and insects are examples of animals without backbones. Fish, birds, and dogs are examples of animals with backbones.

There are many interesting things to learn about animals. Some may even surprise you. For example, the eagle can fly 120 miles per hour, which is very fast, even for an automobile. There are some birds, though, that cannot fly at all. One of these is the penguin. Penguins use their wings for swimming. Penguins can swim as well as a fish can.

There is a crab that climbs palm trees just so it can get coconuts for food. The crab cracks the coconuts with its claws and eats the meat inside.

There is a fish that will drown if it stays underwater too long. It has to come to the top of the water once in a while to gulp air. Sometimes it even crawls out of the water and walks on land. It is called the walking fish!

FORM A

1. Do scientists know how many kinds of animals there are?
2. Do scientists ever discover a new kind of animal?
3. Scientists have divided all the animals into how many main groups?
4. Name an animal given in the passage that does not have a backbone.
5. Do fish have backbones?
6. Can all birds fly?
7. Can the penguin fly?

8. Can a crab climb a tree?
9. What does the crab do with coconuts?
10. How does the walking fish get air to breath?

FORM B

1. About how many kinds of animals do scientists know of?
2. When scientists find that animals are alike, what do they do?
3. Name an animal given in the passage that has a backbone.
4. Do worms have backbones?
5. How fast can the eagle fly?
6. Does the penguin have wings?
7. Can the penguin swim?
8. How does a crab crack coconuts?
9. Can a fish drown?
10. Where does the walking fish walk?

COMPUTERS

What is a computer? It is a machine that can remember everything that it is told. It can solve a problem in less than a second and it doesn't make mistakes like people do. It can work for you and it can help you play. You can play a game, write a song or draw a picture using a computer.

Computers are everywhere. They are in offices, stores, and homes. Sometimes they are in places you cannot see. For example, a microwave oven has a computer inside. When the "start" button is pushed, the computer tells the oven how long to cook and then turns off the oven.

There are computers in many airplanes. The computers tell the pilot how far it is to the next city and how long it will take to get there. Another computer controls the temperature. If the passengers are too warm, the computer turns the heat down.

A computer helps us cross the street. It makes the stoplight turn red, amber, or green.

Did you know that a computer went to the moon? The computer was in a spaceship. Another computer stayed on the earth and talked to the computer in the spaceship.

Computers are used in stores. They count the food that is sold. When only a few boxes of cereal are left, the computer orders more.

Robots have computers inside them. The computer is the brain of the robot. It can make the robot move and talk.

All computers need power to run. Most computers are run by electricity.

FORM A

1. Is a computer a machine?
2. Do computers make mistakes?
3. Name something you can do using a computer.
4. Name two places where there is a computer you can see.
5. What does a computer in a microwave oven tell the oven to do?
6. If the passengers of an airplane are too warm, what does the computer do?
7. How did a computer go to the moon?
8. How is a computer used in a store?
9. The computer is what part of a robot?
10. What do all computers need?

FORM B

1. Do computers forget?
2. About how long does a computer need to solve a problem?
3. Can you always see a computer?
4. Name two places where there is a computer you cannot see.
5. What does the computer in an airplane tell the pilot?
6. How does a computer help us cross the street?
7. Can computers talk to each other?
8. Can a computer count things?
9. What can a computer make a robot do?
10. What kind of power do most computers use?

FOLK TALE

I am going to tell you a folk tale. A folk tale is a story that warns people about being easily fooled. Folk tales often have animals as the main characters. The animals in folk tales can always talk. This tale is called, "Why the Bear is Stumpy-tailed."

One day the Bear met the Fox, who came walking along the road with a string of fish. The Fox had stolen the fish. When the Bear saw that the Fox had a string of fish, he asked, "Where did you get those?"

"I have been out fishing and caught them," lied the Fox.

Now, bears love to eat fish. So the Bear asked the Fox to tell him how to catch fish.

"It is very easy" answered the Fox. "First you must go out on the frozen lake and cut a hole in the ice. Stick your tail down into the water and hold it there as long as you can. Don't worry if your tail freezes a little because that is when the fish bite. The longer you hold it there,

the more fish you'll get. Then pull out your tail with a sideways pull."

Yes, the Bear did just as the Fox had said, and held his tail a long, long time down in the hole, till it was frozen in. Then he pulled it out with a sideways pull and snapped it off. That is why the Bear goes about with a stumpy tail to this very day.

FORM A

1. What kind of story is this?
2. What do folk tales warn people about?
3. What is unusual about the animals in folk tales?
4. What two animals are in this tale?
5. Where did the Fox get the fish?
6. What did the Bear ask the Fox?
7. Did the Fox say it was easy to catch fish?
8. What did the Fox tell the Bear to do with his tail?
9. How did the Fox tell the Bear to pull out his tail?
10. What happened to the Bear's tail?

FORM B

1. What is a folk tale?
2. Who are the main characters in some folk tales?
3. What is the name of this folk tale?
4. What did the Fox have?
5. Where did the Fox say he got the fish?
6. Why did the Bear want fish?
7. Where did the Fox tell the Bear to go?
8. What did the Fox tell the Bear not to worry about?
9. Did the Bear do what the Fox said?
10. Who was fooled in this story?

FOOD

It is fun to compare the foods that people eat in different parts of the world. Imagine that you are traveling around the world and have been invited to eat with friends who live in different countries.

Your first stop is the jungle in Africa. Your friends are having their favorite dinner--roasted monkey. For dessert is a large lump of honeycomb with honey dripping from it. A bee is still in the honeycomb; but it does not sting anyone.

You are off to the North Pole. Today your friends went hunting with spears. They brought back a seal for dinner. They were very hungry so they did not wait to cook it. They just served it raw.

Your next stop is Arabia. You have come a long way and are very thirsty. When you ask for a drink, your friends give you sour milk. It is kept in a leather bag hanging from the ceiling. Your friends like the sour milk. They believe it keeps them strong and well.

Your last stop is China. Tonight's special dinner is bird's-nest soup. The soup is white and has foam on the top. It is made from a swallow's nest. A swallow is a bird that builds its nest high up on steep cliffs. Your friends have climbed the cliffs just to get a swallow's nest for your dinner. They tell you that they don't mind risking a fall for something as delicious as bird's-nest soup.

FORM A

1. Do people eat the same foods all over the world?
2. Are there any people who eat monkeys?
3. What was still in the honeycomb?
4. What do people at the North Pole eat?
5. Why didn't your friends cook the seal?
6. What did your friends in Arabia give you to drink?
7. Why do your friends like sour milk?
8. What was your special dinner in China?
9. What kind of bird builds the special nest used for bird's-nest soup?
10. How did your friends get the nest?

FORM B

1. What was the first country you visited?
2. In Africa, what did you have for dessert?
3. Did the bee sting anyone?
4. What did your friends use to kill the seal?
5. How did your friends serve the seal?
6. In Arabia, where is the milk kept?
7. Where was the last stop in your travels?
8. What color is bird's-nest soup?
9. Where does the swallow build its nest?
10. What might happen to your friends when they go to get the bird's nest?

GLASS

Did you know that nature can make glass? Sometimes during a rain storm, lightning will strike the ground. If the soil happens to be sandy, the lightning will melt some of the sand. When the melted sand cools, it will be a long tube of glass.

The heat from an active volcano can also make glass. The sandy soil around a volcano can become so hot that it melts into a liquid. When the liquid cools, it is a kind of glass.

When people make glass, they mix soda, lime, and sand together and put them into a furnace. These materials are melted by the furnace. When the glass has cooled some, it flows out of the furnace. Then the glass can be rolled, put into a mold, or blown by a glassblower or a machine.

A glassblower dips the end of a blowpipe into hot glass. Some of the glass sticks to the end of the blowpipe. He lifts the blowpipe to his mouth and blows. This makes a bubble of glass. From time to time, the glassblower will reheat the glass to keep it soft. The bubble can then be changed to make many different shapes and sizes of glass objects. All of these objects will be hollow on the inside.

In 1903 a glass-blowing machine was invented. Later another machine was invented just for blowing light bulbs.

FORM A

1. Can only nature make glass?
2. What kind of soil is needed to make glass?
3. What does the glass made by lightning look like?
4. What can a volcano do to sandy soil?
5. Name two materials needed to make glass.
6. When does the glass flow out of the furnace?
7. What kind of pipe does a glassblower use to blow glass?
8. Why does a glassblower reheat the glass?
9. Can a glassblower make a bottle?
10. What was invented in 1903?

FORM B

1. Can lightning make glass?
2. What happens to sand when lightning strikes it?
3. Can a volcano make glass?
4. Can people make glass?
5. Why is a furnace needed to make glass?
6. Can glass be molded?
7. How does a glassblower get melted glass on his blowpipe?
8. When the glassblower makes a bubble, can he change its shape?
9. Can a glassblower make a window?
10. One glass-blowing machine was invented just to make one thing. What was it?

PLANTS

Did you know that there are plants that eat insects? One such plant is called Venus's-flytrap. It has this interesting name because that is what it does--traps flies. Not only does it trap flies and other insects, but it eats them.

Venus's-flytrap grows in the United States. It grows about a foot high with small white flowers at the top. Its leaves have two parts with hinges between them. When an insect lights on a leaf, it closes like a trap and holds the insect inside. After the flytrap has eaten the insect, the trap opens; and the leaves are ready for another victim.

Another meat-eating plant is the pitcher plant. This plant has yellow flowers. Its leaves are shaped like little pitchers. These little pitcher-shaped leaves catch water when it rains. Inside the leaves are little pockets filled with a sweet-smelling juice. Insects land on the leaves to drink this juice. But the leaves are very slippery; and sometimes an insect slides down the leaf and falls into the water. The insect drowns and the pitcher plant has its next meal.

There is one kind of insect that the pitcher plant does not kill. It is the moth. In fact, the moth sometimes makes its home in the pitcher plant.

Most plants get their food from the soil. Venus's-flytrap and the pitcher plant are special plants because they eat meat!

FORM A

1. Can a plant eat an insect?
2. Where does Venus's-flytrap grow?
3. Does Venus's-flytrap have large or small flowers?
4. What happens to an insect that lights on a leaf of Venus's-flytrap?
5. What color is the pitcher plant flower?
6. What shape does a pitcher plant leaf have?
7. What is inside the leaves of a pitcher plant?
8. What makes an insect fall into the water of a pitcher plant?
9. Does the pitcher plant kill all the insects that light on it?
10. Where do most plants get their food?

FORM B

1. What does Venus's-flytrap do?
2. Can a plant eat meat?
3. How tall does Venus's-flytrap grow?
4. What part of Venus's-flytrap closes like a trap?
5. Can a Venus's-flytrap leaf catch more than one insect?
6. What does the pitcher plant do when it rains?
7. Why do insects land on the leaves of pitcher plants?
8. What happens to an insect that falls into the water of a pitcher plant?
9. What insect makes its home in the pitcher plant?

10. What makes Venus's-flytrap and the pitcher plant special?

SHELLS

Many kinds of animals have shells. Turtles and snails are two kinds of animals with shells. They grow their shells on the outside of their bodies.

Some animals grow shells on the inside of their bodies. For example, the squid has a shell just under its skin. The squid's shell makes its body strong.

Many other kinds of animals have shells too, but they only use their shells at the beginning of their lives. Birds grow inside an egg shell. But, when they hatch out of the shell, they leave it behind.

Shells can be very big. The biggest shell belongs to the giant clam. The giant clam can grow a shell that is four feet long. A small person could climb inside a giant clam shell!

Shells can also be very tiny. There is one kind of shell that is no bigger than a grain of sand.

Seashells can be used to make a kind of cloth. This cloth has such a shiny surface that it looks like it is made from gold. It is fine and soft and very beautiful.

Another beautiful thing that comes from seashells is the pearl. Pearls are sometimes formed inside an oyster. When the oyster is broken open, the pearl may be taken out and used to make jewelry. Today there are pearl farmers who raise thousands of oysters so that they can sell the pearls that are inside.

FORM A

1. Do all animals have shells?
2. Do any animals grow shells on the inside of their bodies?
3. How is the squid helped by his shell?
4. What kind of shell do birds have?
5. How long is the giant clam shell?
6. Name something beautiful that comes from seashells.
7. What does cloth made from seashells look like?
8. Where do pearls come from?
9. What can pearls be used for?
10. Why do pearl farmers raise oysters?

FORM B

1. Name an animal with a shell on the outside of his body.
2. Where is the squid's shell?
3. What animal only uses its shell at the beginning of its life?
4. What animal has the biggest shell?
5. Can a shell be as small as a grain of sand?
6. Can seashells be used to make cloth?
7. How could you describe seashell cloth?
8. How could you get a pearl out of an oyster?
9. Who raises oysters?
10. What do pearl farmers do with the pearls?

STARS

The biggest thing that anyone knows of is a star. A star is a huge ball of burning gas in the sky. When we look at the sky at night, we can see about two thousand stars. These are the very brightest stars. There are many more stars we cannot see because they are too far away.

People who study the stars are called astronomers. Astronomers use a telescope to look at the stars. The telescope makes the stars look much bigger. With a telescope, astronomers can see about three billion stars.

Astronomers measure the color of stars. They have found that if a star gives off blue light it is extremely hot. If it looks red, it is a cooler star. A yellow star has a temperature somewhere between the blue and red stars.

Astronomers also measure the size of stars. They have named the biggest stars supergiants. Our sun is very small when compared to a supergiant star.

Astronomers have discovered that there are groups of stars. They have named one of these groups the Milky Way. The earth, the sun, and all the stars that we can see without a telescope are part of the Milky Way. Astronomers tell us that the Milky Way looks something like a big pancake in the sky.

Did you know that our sun is a star? It doesn't look like a star because it is so close to the earth. It is just close enough to give us light and keep us warm.

FORM A

1. What is the biggest thing that anyone knows about?
2. About how many thousand stars can we see at night?
3. Why can't we see all the stars?
4. What do astronomers study?
5. What are telescopes used for?

6. What do astronomers measure about stars?
7. What color is the hottest star?
8. Which is bigger, a supergiant or the sun?
9. What is the Milky Way?
10. What do we call the star that is closest to the earth?

FORM B

1. What is a star made of?
2. Can we see all the stars?
3. What does a telescope do?
4. With a telescope, about how many billion stars can astronomers see?
5. What three colors can the stars be?
6. What are the biggest stars called?
7. What does the Milky Way look like?
8. Is the earth in the Milky Way?
9. Why doesn't our sun look like a star?
10. What does the sun give us?

THE WHEEL

Many years ago people learned that animals could help them carry things. In different parts of the world, the ox, donkey, water buffalo, horse and camel were used to carry things. These animals were called "beasts of burden."

Sometimes the things that people wanted to carry were too big to put on the back of an animal. So, a sled was made to be loaded with things and pulled by the animal. This worked fine on the snow because the snow made the ground slick and smooth. The sled did not work so well on regular ground because it was very hard for the animal to pull.

Finally someone thought of cutting a slice from the end of a log and making a hole in its center. This was the first wheel. When two wheels were put on the ends of a pole and fastened to the sled, the first cart was invented. This worked better than the sled because it was much easier to pull.

Over many years, people made the wheel bigger and better. They cut away part of the wood so the wheel would be lighter. They also covered the rim of the wheel with copper so it would not wear out so quickly. People also learned to make roads. Roads made it easier to go from one place to another. The early roads were paved with stones.

In 1869 rubber tires were invented. The first bicycles and cars used this kind of wheel. Today, we still use rubber tires; but we fill them with air.

FORM A

1. Many years ago how did people carry things?
2. Is a dog a "beast of burden"?
3. How were very big things carried?
4. Why was a sled easier to pull in the snow?
5. What was the first wheel made from?
6. What was put between the two wheels?
7. Why was a cart better than a sled?
8. What did people put around the rim of the wheel?
9. What were the first roads covered with?
10. What kind of wheel did the first bicycles and cars use?

FORM B

1. Name two "beasts of burden."
2. How did the animals move a sled?
3. When was it very hard for the animal to pull the sled?
4. What was put in the center of the wheel?
5. What did two wheels, a pole, and a sled make?
6. Why did people cut away part of the wood on the wheel?
7. Besides the wheel, what else did people have to make it easier to go from one place to another?
8. What was invented in 1869?
9. Do we use rubber tires today?
10. What do today's rubber tires have inside them?

APPENDIX B

EVALUATOR LETTER

EVALUATION FORM

Central State University
100 North University Drive • Edmond, Oklahoma 73034 • 405-341-2980

College of Education
Department of Psychology and Personnel Services

Dear Evaluator:

Thank you for evaluating the enclosed test materials.

These passages and questions were written to serve as the material for a listening comprehension test. Listening comprehension, for the purposes of this test, is defined as the amount of information correctly recalled after a verbal presentation of the passages.

The final form of the test will contain five passages, each followed by ten questions. The passages and questions will be read orally and the examinees will answer the questions orally (if in kindergarten, first, or second grade) or by writing (if in third, fourth, fifth, or sixth grade).

Ten passages, each with twenty questions, were written with the following criteria in mind: (1) The passages contain only material that the examinees would not be likely to know without having heard the passages. (2) The passages and questions contain typical oral language familiar to school children. (3) The passage material will be selected to minimize bias (age, gender, cultural, ethnic, and racial). (4) The questions will call for simple and minimal responses.

Please evaluate these materials by responding to the questions on the enclosed Evaluation Form, then return the Evaluation Form and the test materials in the envelope provided.

Thank you,

Peggy Kerr

Peggy Kerr, Coordinator
School Psychometry
and School Psychology

EVALUATION FORM

Evaluator's Name_____

Title/Position_____

Address_____

Phone_____

1. Ranking Passages: Please rank the passages with 1 being the best passage for measuring listening comprehension and 10 being the worst.

Passage Name	Rank
ALPHABET	_____
ANIMALS	_____
COMPUTERS	_____
FOLK TALE	_____
FOOD	_____
GLASS	_____
PLANTS	_____
SHELLS	_____
STARS	_____
THE WHEEL	_____

Comments_____

_____

_____

2. Does this test, overall, appear to measure listening comprehension?_____If no, why not?

_____

_____

APPENDIX C

PARENT PERMISSION LETTER

DIRECTIONS FOR ADMINISTERING AND SCORING
THE LISTENING COMPREHENSION TEST
FIELD STUDY

LISTENING COMPREHENSION TEST PASSAGES

PROTOCOLS

Central State University
100 North University Drive • Edmond, Oklahoma 73034 • 405-341-2980

College of Education
Department of Psychology and Personnel Services

Dear Parent,

I am field testing a listening comprehension exercise at your child's school.

This 30 minute exercise involves the following steps: Five short passages are read aloud. Following each passage, ten questions are asked about the passsge. Children in kindergarten through second grade respond verbally, while children in third through sixth grade respond by writing. The passages consists of subjects familiar to children, such as plants and animals.

The results of this field study will be used to develop a standardized listening comprehension test.

Please give permission for your child to participate in this exercise by signing and returning this letter.

Thank you for your cooperation,

Peggy Kerr

Peggy Kerr, Coordinator
Psychometry and School Psychology

I give permission for my child,_____, to participate in this field study.

_____
Parent

DIRECTIONS FOR ADMINISTERING AND SCORING
THE LISTENING COMPREHENSION TEST
FIELD STUDY

Test Description

The field test consists of two forms, A and B. An examinee will take only one of the forms. Each form contains five passages (these are the same for both forms) and 50 questions. Both the passages and questions are read aloud to the examinee. The test may be administered to children from 5 to 12 years of age. Administration time is 20 to 30 minutes. The field test may be individually administered to any examinee or it may be group administered to children in the third grade or above.

Test Contents

1. One set of five passages for forms A and B.
2. Protocol for group administration of forms A or B.
3. Protocol for individual administration of form A. (This protocol is also used as the source for the questions and scoring of the group administered protocol, form A.)
4. Protocol for individual administration of form B. (This protocol is also used as the source for the questions and scoring of the group administered protocol, form B.)

Directions for Group Administration, Forms A & B

Give each examinee a "protocol for group administration". Instruct the examinees to complete the identifying information on their protocols and to circle the A or B for "Form" as you so indicate.

Say: **"I am going to read some stories and after each story, ask ten questions. It is important that you listen carefully, because I cannot repeat the stories. I will repeat each question only one time. Write your answer on the line that has the same number as the question. If you do not know the answer, draw a line in the blank for that question."**

Before reading each passage, say, **" Please listen carefully."** Then read the passage at a normal speaking rate. Using the "protocol for individual administration, form A or B," read each question for that passage and repeat it one time. Allow 30 seconds per question for examinees to write their answers.

Scoring

Correct responses are listed on the "protocol for individual administration, form A or B." Circle "P" if one of these responses is given by the examinee. Spelling errors do not make a response wrong. If the examinee gives a response that is not listed, circle "F." Count the total number of "P" scored responses and write on the "Total" line at the end of the protocol.

Directions for Individual Administration, Forms A & B

Use the "protocol for individual administration," form A or B. Complete the identifying information for the examinee. Do not allow the examinee to see the protocol before or during administration because correct responses are listed on the protocol.

Say: **"I am going to read some stories and after each story, ask ten questions. It is important that you listen carefully, because I cannot repeat stories. I will repeat a question if you need to hear it again."**

Before reading each passage, say, **"Please listen carefully."** Then read the passage at a normal speaking rate. Read each question for that passage. Allow a maximum of 30 seconds per question for the examinee to respond. Record the exact response of the examinee. If a response is unclear, do not query.

Scoring

Correct responses are listed on the protocol. Circle "P" if one of these responses is given by the examinee. If the examinee gives a response that is not listed on the protocol, circle "F." Count the total number of "P" scored responses and write on the "Total" line at the end of the protocol.

LISTENING COMPREHENSION TEST PASSAGES

ALPHABET

A long time ago people did not have an alphabet. They just drew little pictures when they wanted to write something. Each picture stood for a word. This did not work very well because they needed many pictures just to write a sentence.

The first people to use an alphabet were the Semites. The Semites lived about 3,500 years ago. They decided to write their language using signs. Their signs were like simple little pictures with one big difference. The signs did not stand for words. Each sign stood for a sound. For example, a picture of an ox head stood for the sound of "a." A house stood for the sound of "b." Waves of water was the sign for the "m" sound. The Semites discovered that they only needed about 30 signs or letters to write their language.

Other groups of people began to use the Semite alphabet. The Greeks and Romans changed the shapes of some of the letters. They gave up a few of the letters they did not need. They also added one or two letters of their own.

Today most languages are written with alphabets made from the one invented by the Semites long ago. A few languages are still written with signs that are not letters of an alphabet. Chinese writing, for instance, is done with such signs. But writing with an alphabet is much more common.

ANIMALS

Did you know that the world has so many animals that no one knows how many kinds there are? Scientists have found almost a million kinds of animals so far. Every year, hundreds of new kinds of animals are discovered.

Scientists have studied how animals are alike and how they are different. When they find that animals are alike, they put them into a group. They have divided all the animals into two main groups--animals without backbones and animals with backbones. Worms, starfish, and insects are examples of animals without backbones. Fish, birds, and dogs are examples of animals with backbones.

There are many interesting things to learn about animals. Some may even surprise you. For example, the

eagle can fly 120 miles per hour, which is very fast, even for an automobile. There are some birds, though, that cannot fly at all. One of these is the penguin. Penguins use their wings for swimming. Penguins can swim as well as a fish can.

There is a crab that climbs palm trees just so it can get coconuts for food. The crab cracks the coconuts with its claws and eats the meat inside.

There is a fish that will drown if it stays underwater too long. It has to come to the top of the water once in a while to gulp air. Sometimes it even crawls out of the water and walks on land. It is called the walking fish!

FOLK TALE

I am going to tell you a folk tale. A folk tale is a story that warns people about being easily fooled. Folk tales often have animals as the main characters. The animals in folk tales can always talk. This tale is called, "Why the Bear is Stumpy-tailed."

One day the Bear met the Fox, who came walking along the road with a string of fish. The Fox had stolen the fish. When the Bear saw that the Fox had a string of fish, he asked, "Where did you get those?"

"I have been out fishing and caught them," lied the Fox.

Now, bears love to eat fish. So the Bear asked the Fox to tell him how to catch fish.

"It is very easy" answered the Fox. "First you must go out on the frozen lake and cut a hole in the ice. Stick your tail down into the water and hold it there as long as you can. Don't worry if your tail freezes a little because that is when the fish bite. The longer you hold it there, the more fish you'll get. Then pull out your tail with a sideways pull."

Yes, the Bear did just as the Fox had said, and held his tail a long, long time down in the hole, till it was frozen in. Then he pulled it out with a sideways pull and snapped it off. That is why the Bear goes about with a stumpy tail to this very day.

FOOD

It is fun to compare the foods that people eat in different parts of the world. Imagine that you are traveling around the world and have been invited to eat with friends who live in different countries.

Your first stop is the jungle in Africa. Your friends are having their favorite dinner--roasted monkey. For dessert is a large lump of honeycomb with honey dripping from it. A bee is still in the honeycomb; but it does not sting anyone.

You are off to the North Pole. Today your friends went hunting with spears. They brought back a seal for dinner. They were very hungry so they did not wait to cook it. They just served it raw.

Your next stop is Arabia. You have come a long way and are very thirsty. When you ask for a drink, your friends give you sour milk. It is kept in a leather bag hanging from the ceiling. Your friends like the sour milk. They believe it keeps them strong and well.

Your last stop is China. Tonight's special dinner is bird's-nest soup. The soup is white and has foam on the top. It is made from a swallow's nest. A swallow is a bird that builds its nest high up on steep cliffs. Your friends have climbed the cliffs just to get a swallow's nest for your dinner. They tell you that they don't mind risking a fall for something as delicious as bird's-nest soup.

PLANTS

Did you know that there are plants that eat insects? One such plant is called Venus's-flytrap. It has this interesting name because that is what it does--traps flies. Not only does it trap flies and other insects, but it eats them.

Venus's-flytrap grows in the United States. It grows about a foot high with small white flowers at the top. Its leaves have two parts with hinges between them. When an insect lights on a leaf, it closes like a trap and holds the insect inside. After the flytrap has eaten the insect, the trap opens; and the leaves are ready for another victim.

Another meat-eating plant is the pitcher plant. This plant has yellow flowers. Its leaves are shaped like little pitchers. These little pitcher-shaped leaves catch water when it rains. Inside the leaves are little pockets filled

with a sweet-smelling juice. Insects land on the leaves to drink this juice. But the leaves are very slippery; and sometimes an insect slides down the leaf and falls into the water. The insect drowns and the pitcher plant has its next meal.

There is one kind of insect that the pitcher plant does not kill. It is the moth. In fact, the moth sometimes makes its home in the pitcher plant.

Most plants get their food from the soil. Venus's-flytrap and the pitcher plant are special plants because they eat meat!

PROTOCOL
LISTENING COMPREHENSION FIELD TEST
GROUP ADMINISTRATION

Name_____

School_____

Circle one in each category below.

Age 7 8 9 10 11 12 Gender Male Female

Grade 3rd 4th 5th 6th Form A B

Instructions

Listen carefully to each story. After each story, you will be asked ten questions. Write your answers on the lines below. If you do not know the answer, draw a line in the blank.

ALPHABET

P/F 1._____

P/F 2._____

P/F 3._____

P/F 4._____

P/F 5._____

P/F 6._____

P/F 7._____

P/F 8._____

P/F 9._____

P/F 10._____

ANIMALS

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

FOLK TALE

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

FOOD

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

PLANTS

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

_____ TOTAL P

PROTOCOL
LISTENING COMPREHENSION FIELD TEST
INDIVIDUAL ADMINISTRATION

FORM A

Examinee_____

Examiner_____

Circle one in each category below

Age 5 6 7 8 9 10 11 12 Gender Male Female

Grade K 1st 2nd 3rd 4th 5th 6th

ALPHABET

P/F 1. Have people always had an alphabet? (no)_____

P/F 2. In writing with pictures, what does each picture stand for? (a word)_____

P/F 3. Who were the first people to use an alphabet? (Semites)_____

P/F 4. How long ago did the Semites live? (3,500 years)

_____

P/F 5. How did the Semites write the sound of "a"? (ox/ox head)_____

P/F 6. Waves of water was the sign for what sound? (m)___

P/F 7. Was the Semite alphabet ever changed? (yes)_____

P/F 8. Who changed the shape of the letters? (Greeks and/or Romans)_____

P/F 9. Today, do people use alphabets made from the one invented by the Semites? (yes)_____

P/F 10. Does Chinese writing use an alphabet? (no)_____

ANIMALS

P/F 1. Do scientists know how many kinds of animals there are? (no)_____

P/F 2. Do scientists ever discover a new kind of animal? (yes)_____

P/F 3. Scientists have divided all the animals into how many main groups? (two)_____

P/F 4. Name an animal given in the passage that does not have a backbone. (worm/starfish/insect)_____

P/F 5. Do fish have backbones? (yes)_____

P/F 6. Can all birds fly? (no)_____

P/F 7. Can a penguin fly? (no)_____

P/F 8. Can a crab climb a tree? (yes)_____

P/F 9. What does a crab do with coconuts? (cracks them/ eats them)_____

P/F 10. How does the walking fish get air to breath? (comes to the top of the water)_____

FOLK TALE

P/F 1. What kind of story is this? (folk tale)_____

P/F 2. What do folk tales warn people about? (being fooled)_____

P/F 3. What is unusual about the animals in folk tales? (They talk.)_____

P/F 4. What two animals are in this tale? (bear and fox)

_____

P/F 5. Where did the Fox get the fish? (stole them)_____

P/F 6. What did the Bear ask the Fox? (Where did you get those? /to tell him how to catch fish)_____

P/F 7. Did the Fox say it was easy to catch fish? (yes)___

P/F 8. What did the Fox tell the Bear to do with his tail? (stick it into the water)_____

P/F 9. How did the Fox tell the Bear to pull out his tail? (with a sideways pull/sideways)_____

P/F 10. What happened to the Bear's tail? (It snapped off)

_____

FOOD

P/F 1. Do people all over the world eat the same foods? (no)_____

P/F 2. Are there any people who eat monkeys? (yes)_____

P/F 3. What was still in the honeycomb? (a bee/honey)_____

P/F 4. What do people at the North Pole eat? (seal)_____

P/F 5. Why didn't your friends cook the seal? (They were very hungry/too hungry.)_____

P/F 6. What did your friends in Arabia give you to drink? (sour milk)_____

P/F 7. Why do your friends like sour milk? (Response containing the words, "strong" and/or "well.")

_____

P/F 8. What was your special dinner in China? (bird's-nest soup)_____

P/F 9. What kind of bird builds the special nest used for bird's-nest soup? (swallow)_____

P/F 10. How did your friends get the nest? (They climbed the cliffs/climbed.)_____

PLANTS

P/F 1. Can a plant eat an insect? (yes)_____

P/F 2. Where does Venus's-flytrap grow? (U.S.)_____

P/F 3. Does Venus's-flytrap have large or small flowers? (small)_____

P/F 4. What happens to an insect that lights on a leaf of Venus's-flytrap? (It is trapped/it is eaten)

_____

P/F 5. What color is the pitcher plant flower? (yellow)

_____

P/F 6. What shape does a pitcher plant leaf have? (pitcher shaped)_____

P/F 7. What is inside the leaves of a pitcher plant? (juice/sweet-smelling juice)_____

P/F 8. What makes an insect fall into the water of a pitcher plant? (slippery leaves)_____

P/F 9. Does the pitcher plant kill all the insects that light on it? (no)_____

P/F 10. Where do most plants get their food? (from the soil/soil)_____

_____Total P

PROTOCOL
LISTENING COMPREHENSION FIELD TEST
INDIVIDUAL ADMINISTRATION

FORM B

Examinee_____

Examiner_____

Circle one in each category below.

Age 5 6 7 8 9 10 11 12 Gender Male Female

Grade K 1st 2nd 3rd 4th 5th 6th

ALPHABET

P/F 1. A long time ago, what did people do when they wanted to write something? (drew pictures)_____

P/F 2. Writing with pictures does not work very well. Why? (need many pictures)_____

P/F 3. What did the Semites invent? (alphabet)_____

P/F 4. In the Semite alphabet, what did the signs stand for? (sounds)_____

P/F 5. How did the Semites write the sound of "b"? (house)_____

P/F 6. How many letters were in the Semite alphabet? (30)_____

P/F 7. What did the Greeks and Romans do to the Semite alphabet? (used it/changed it/added letters/gave up letters)_____

P/F 8. Were any letters ever added to the alphabet? (yes)_____

P/F 9. Are any languages written with signs that are not letters of an alphabet? (yes)_____

P/F 10. What is the most common way to write a language? (with an alphabet/letters)_____

ANIMALS

P/F 1. About how many kinds of animals do scientists know of? (million)_____

P/F 2. When scientists find that animals are alike, what do they do? (put them into a group)_____

P/F 3. Name an animal given in the passage that has a backbone. (fish/bird/dog)_____

P/F 4. Do worms have backbones? (no)_____

P/F 5. How fast can the eagle fly? (120 mph)_____

P/F 6. Does the penguin have wings? (yes)_____

P/F 7. Can the penguin swim? (yes)_____

P/F 8. How does a crab crack coconuts? (with its claws)___

_____

P/F 9. Can a fish drown? (yes)_____

P/F 10. Where does the walking fish walk? (on land)_____

FOLK TALE

P/F 1. What is a folk tale? (a response containing the word, "story")_____

P/F 2. Who are the main characters in some folk tales? (animals)_____

P/F 3. What is the name of this folk tale? (Why the Bear is Stumpy-tailed/ response containing the words, "bear" and "tail")_____

P/F 4. What did the Fox have? (fish)_____

P/F 5. Where did the Fox say he got the fish? (had been fishing/caught them)_____

P/F 6. Why did the Bear want fish? (to eat/bears love fish)_____

P/F 7. Where did the Fox tell the Bear to go? (to the lake)_____

P/F 8. What did the Fox tell the Bear not to worry about? (tail freezing)_____

P/F 9. Did the Bear do what the Fox said? (yes)_____

P/F 10. Who was fooled in this story? (the Bear)_____

FOOD

P/F 1. What was the first country you visited? (Africa)_____

P/F 2. In Africa, what did you have for dessert? (honey/ honeycomb)_____

P/F 3. Did the bee sting anyone? (no)_____

P/F 4. What did your friends use to kill the seal? (spear)_____

P/F 5. How did your friends serve the seal? (raw)_____

P/F 6. In Arabia, where is the milk kept? (bag)_____

P/F 7. Where was the last stop in your travels? (China)_____

P/F 8. What color is bird's-nest soup? (white)_____

P/F 9. Where does the swallow build its nest? (cliffs)_____

P/F 10. What might happen to your friends when they go to get the bird's nest? (a fall)_____

PLANTS

P/F 1. What does Venus's-flytrap do? (It traps flies/eats insects)_____

P/F 2. Can a plant eat meat? (yes)_____

P/F 3. How tall does Venus's-flytrap grow? (a foot)_____

P/F 4. What part of Venus's-flytrap closes like a trap? (leaves)_____

P/F 5. Can a Venus's-flytrap leaf catch more than one insect? (yes)_____

P/F 6. What does the pitcher plant do when it rains? (catches water/catches rain)_____

P/F 7. Why do insects land on the leaves of pitcher plants? (to drink the juice/to drink_____

P/F 8. What happens to an insect that falls into the water of a pitcher plant? (It drowns/it dies/it is eaten)_____

P/F 9. What insect makes its home in the pitcher plant? (moth)_____

P/F 10. What makes Venus's-flytrap and the pitcher plant special? (They eat meat/they eat insects)_____

_____Total P

APPENDIX D

PARENT PERMISSION LETTER

DIRECTIONS FOR ADMINISTERING AND SCORING THE LISTENING COMPREHENSION TEST

PROTOCOLS

Central State University
100 North University Drive • Edmond, Oklahoma 73034 • 405-341-2980

College of Education
Department of Psychology and Personnel Services

Dear Parent:

I am conducting a listening comprehension study at your child's school.

This study involves the following steps: Five short passages are read aloud. Following each passage, questions are asked about the passage. Children in kindergarten through second grade respond verbally, while children in third through sixth grade respond by writing. The passages consist of subjects familiar to children, such as plants and animals. These steps will be repeated with the same students one week later. Each session is approximately 30 minutes in length.

The results of this study will be used to develop a standardized listening comprehension test.

Please give permission for your child to participate in this study by signing and returning this letter.

Thank you for your cooperation,

Peggy Kerr

Peggy Kerr, Coordinator
Psychometry and School Psychology

I give permission for my child,_____ to participate in this study.

_____
Parent

DIRECTIONS FOR ADMINISTERING AND SCORING
THE LISTENING COMPREHENSION TEST

Test Description

This listening comprehension test contains five passages and 50 questions. Both the passages and questions are read aloud to the examinee. The test may be administered to children from 5 to 12 years of age. Administration time is 20 to 30 minutes. The test may be administered individually to any examinee. It may be group administered only to children who are in the third grade or above.

Test Contents

1. Set of five passages
2. Protocol for group administration
3. Protocol for individual administration (This protocol is also used as the source for the questions when administering the test to a group.)

Directions for Group Administration

Give each examinee a "protocol for group administration." Instruct the examinees to complete the identifying information on their protocol. Say: **"I am going to read some stories and then ask some questions about the stories. It is important to listen carefully, because I cannot repeat the stories. I will repeat each question only one time. Write your answer on the line that has the same number as the question. If you do not know the answer, draw a line in the blank for that question."**

Before reading each passage, say, **"Please listen carefully."** Then read the passage at a normal speaking rate. When you finish reading the passage, use the "protocol for individual administration," and read each question for the passage, repeating it one time. Allow 30 seconds per question for examinees to write their answers.

Directions for Individual Administration

Use the "protocol for individual administration." Complete the identifying information for the examinee.

Say: **"I am going to read some stories and then ask some questions about the stories. It is important to listen carefully, because I cannot repeat the stories. I will repeat a question if you need to hear it again."**

Before reading each passage, say, **"Please listen carefully."** Then read the passage at a normal speaking rate. After reading the passage, read each question for the

passage. Allow a maximum of 30 seconds per question for the examinee to respond. Record the exact response of the examinee. If a response is unclear, do not query.

Scoring Guide

Guidelines for scoring and examples of "pass" and "fail" responses are listed below. Circle "P" on the protocol if the examinee's response meets the guideline. If the examinee gives a response that does not meet the guideline, circle "F." Count the total number of "P" scored responses and write on the "Total" line at the end of the protocol.

On the "protocol for group administration," spelling errors do not make a response wrong.

ALPHABET

1. A long time ago, what did people do when they wanted to write something?

Response must contain the word, "picture/s."

"Pass"	"Fail"
drew pictures	drew it
used pictures	used signs
made pictures	alphabet
	(synonym for "pictures")

2. Writing with pictures does not work very well. Why?

Response must indicate that many pictures were needed.

"Pass"	"Fail"
need many pictures	it took a long time
too many	they didn't understand
it took a lot of them	

3. Who were the first people to use an alphabet?

"Pass"	"Fail"
Semite/s	(all other responses)

4. In the Semite alphabet, what did the signs stand for?

"Pass"	"Fail"
sound/s	(all other responses)

5. How did the Semites write the sound of "a"?

Response must contain the word, "ox." Some children may draw their answer during group administration. This is scored as a fail.

"Pass"	"Fail"
ox	fox
drew an ox	cow
They made an ox head	buffalo
head of an ox	(synonym for "ox")

6. How did the Semites write the sound of "b"?

Response must contain the word, "house." Some children may draw their answer during group administration. This is scored as a fail.

"Pass"	"Fail"
house	(synonym for "house")
drew a house	
made a house	

7. How many letters were in the Semite alphabet?

During group administration, response may be written in words or numerals.

"Pass"	"Fail"
30	(all other responses)
about thirty	

8. What did the Greeks and Romans do to the Semite alphabet?

Response must contain a verb which indicates (1) use of the alphabet **or** (2) some change in the shape or number of letters.

"Pass"	"Fail"
used it	more
changed the shape	made it into letters
added letters	made it into words
made more	made their own
gave up some	
took some out	

9. What is the most common way to write a language?

Response must contain the word (1) "alphabet" **or** (2) "letters."

"Pass"	"Fail"
with an alphabet	with signs
in letters	pictures
	abc's

ANIMALS

1. About how many kinds of animals do scientists know of?

Response must contain the word, "million." If million is written in numbers during group administration, it must contain the correct number of zeros.

"Pass"	"Fail"
1,000,000	millions
million	
about a million	
almost a million	

2. When scientists find that animals are alike, what do they do?

Response must contain the concept of grouping.

"Pass"	"Fail"
group them	count them
Put them into groups	study them
categorize them	
classify them	
divided them into groups	
groups with and without backbones	

3. Scientists have divided all the animals into how many main groups?

Response may be written as a word or numeral during group administration.

"Pass"	"Fail"
2	(all other responses)
two	

4. Name an animal given in the story that does not have a backbone.

"Pass"	"Fail"
worm	(all other responses)
starfish	
insect	

5. Name an animal given in the story that has a backbone.

"Pass"	"Fail"
fish	(all other responses)
bird	
dog	

6. Do fish have backbones?

"Pass"	"Fail"
yes	(all other responses)
(an affirmative response)	

7. How fast can the eagle fly?

"Pass"	"Fail"
120 ("mph" is optional)	(all other responses)

8. Can a fish drown?

"Pass"	"Fail"
yes	(all other responses)
some can	
(an affirmative response)	

9. How does the walking fish get air to breath?

Response must indicate that he breaths at the top of the water.

"Pass"	"Fail"
comes to the top of the water	crawls out on land
comes to the surface	comes out of the water
goes up	walks on land
	gulps air

FOLK TALE

1. What is a folk tale?

Response must contain the word, "story."

"Pass"	"Fail"
story about animals	fairy tale
story that warns people	a tale that isn't true
not a true story	it is a lie
	(synonym for "story")

2. Who are the main characters in some folk tales?

Response must contain the word, "animal/s."

"Pass"	"Fail"
animals that talk	Bear and Fox

3. What is unusual about the animals in folk tales?

Response must contain the word, "talk."

"Pass"	"Fail"
they talk	they are fooled
talking animals	(synonym for "talk")

4. What is the name of this folk tale?

Response must contain the words (1) "bear" **and** (2) "tail."
During group administration, "tail" may be spelled "tale."

"Pass"	"Fail"
Why the Bear is Stumpy-tailed	The Bear and the Fox
Why the Bear is Stubby-tailed	
The Bear's Tail	

5. Where did the Fox get the fish?

Response must contain the concept of stealing.

"Pass"	"Fail"
he stole them	from the bear
they were stolen	from the lake

6. What did the Fox tell the Bear not to worry about?

Response must contain the words (1) "tail" **and** (2) a form of the verb, "freeze."

"Pass"	"Fail"
his tail freezing	his tail getting cold
his tail getting frozen	his tail

7. How did the Fox tell the Bear to pull out his tail?

Response must contain the word, "side" or "sideways."

"Pass"	"Fail"
with a sideways pull	with a jerk
to the side	

8. Who was fooled in this story?

"Pass"	"Fail"
Bear	(all other responses)

FOOD

1. What was the first country you visited?

Response must contain the word, "Africa."

"Pass"	"Fail"
Africa	jungle
jungles of Africa	

2. Are there any people who eat monkeys?

"Pass"	"Fail"
yes	(all other responses)
Africans	
(an affirmative response)	

3. What do people at the North Pole eat?

"Pass"	"Fail"
seal/s	frozen food
	penguins

4. What did your friends use to kill the seal?

"Pass"	"Fail"
spear/s	(synonym for "spear")

5. Why didn't your friends cook the seal?

Response must indicate that they didn't wait because they were so hungry.

"Pass"	"Fail"
they were very hungry	they liked it raw
too hungry	
can't wait	

6. In Arabia, where is the milk kept?

Response must contain the word, "bag."

"Pass"	"Fail"
leather bag	on the ceiling
bag hanging from the ceiling	(synonym for "bag")

7. Why do your friends like sour milk?

Response must contain one of the words (1) "strong" **or** (2) "well."

"Pass"	"Fail"
it makes them strong	they like it
they think it keeps them well	it keeps them healthy
they believe it keeps them strong and well	

8. What was your special dinner in China?

Response must contain the word, "nest."

"Pass"	"Fail"
bird's-nest	bird soup
bird's-nest soup	swallow soup
nest soup	
swallow's nest	

9. What color is bird's-nest soup?

"Pass"	"Fail"
white	(all other responses)

10. Where does the swallow build its nest?

Response must contain the word, "cliff/s."

"Pass"	"Fail"
cliffs	mountains
in the cliffs	trees
on high cliffs	(synonym for "cliff")

11. How did your friends get the nest?

Response must contain the verb, "climb." If the examinee used another (incorrect) noun, such as "mountain" or "tree" in item 10 and used that noun again in this response with the verb, "climb," the response is scored as a pass in item 11.

"Pass"	"Fail"
they climbed	went up the cliff
climbed the cliffs	(synonym for "climb")

12. What might happen to your friends when they go to get the bird's-nest?

Response must contain the word, "fall."

"Pass"	"Fail"
fall	killed
a fall	hurt
they might fall	they could die

PLANTS

1. Can a plant eat meat?

"Pass"	"Fail"
yes (an affirmative response)	(all other responses)

2. Where does Venus's-flytrap grow?

Response must contain the words, "U.S."

"Pass"	"Fail"
United States	South America
South part of U.S.	South

3. How tall does Venus's-flytrap grow?

Response must contain the word, "foot."

"Pass"	"Fail"
a foot	(synonym for foot)
one foot	
about a foot high	

4. What part of Venus's-flytrap closes like a trap?

Response must contain the word, "leaf."

"Pass"	"Fail"
leaf	petals
leaves	flower

5. What shape does a pitcher plant leaf have?

Response must contain the word, "pitcher." "Pitcher" may be spelled "picture" during group administration.

"Pass"	"Fail"
pitcher	(all other responses)
pitcher shaped	

6. What is inside the leaves of a pitcher plant?

Response must contain the word, "juice."

"Pass"	"Fail"
sweet-smelling juice	nectar
sweet juice	water
	(synonym for "juice")

7. Why do insects land on the leaves of pitcher plants?

Response must contain the verb, "drink." If the examinee used another noun in item 6, such as "water," and used that noun again in this response with the verb, "drink," the response is scored as a pass in item 7.

"Pass"	"Fail"
to drink	to get the juice
to drink the juice	to eat
	(synonym for "drink")

8. What causes an insect to fall into the water of a pitcher plant?

Response must contain the idea of slipping or sliding.

"Pass"	"Fail"
slippery leaves	the juice
he slips	the leaves
the leaves are slick	
he slides down	

9. What happens to an insect that falls into the water of a pitcher plant?

Response must indicate that the insect (1) dies **or** (2) becomes the food of the plant.

"Pass"	"Fail"
it drowns	it drinks the water
it dies	it lives there
it is eaten	
it is the next meal	

10. What insect makes its home in the pitcher plant?

"Pass"	"Fail"
moth	(all other responses)

11. Where do most plants get their food?

Response must contain the word, "soil."

"Pass"	"Fail"
soil	from insects
from the soil	from water
	(synonym for "soil")

12. What makes Venus's-flytrap and the pitcher plant special?

Response must contain the verb, "eat."

"Pass"	"Fail"
they eat meat	they trap flies
they eat insects	
they are meat-eating plants	

PROTOCOL
LISTENING COMPREHENSION TEST
GROUP ADMINISTRATION

Name_____

School_____

Circle one in each category below.

Age 7 8 9 10 11 12 Gender Male Female

Grade 3rd 4th 5th 6th

Instructions

Listen carefully to each story. After each story, you will be asked some questions. Write your answers on the lines below. If you do not know the answer, draw a line in the blank.

ALPHABET

P/F 1._____

P/F 2._____

P/F 3._____

P/F 4._____

P/F 5._____

P/F 6._____

P/F 7._____

P/F 8._____

P/F 9._____

ANIMALS

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

FOLK TALE

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

FOOD

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

P/F 11. _____

P/F 12. _____

PLANTS

P/F 1. _____

P/F 2. _____

P/F 3. _____

P/F 4. _____

P/F 5. _____

P/F 6. _____

P/F 7. _____

P/F 8. _____

P/F 9. _____

P/F 10. _____

P/F 11. _____

P/F 12. _____

_____ TOTAL P

PROTOCOL
LISTENING COMPREHENSION TEST
INDIVIDUAL ADMINISTRATION

Examinee_____

Examiner_____

Circle one in each category below.

Age 5 6 7 8 9 10 11 12 Gender Male Female

Grade K 1st 2nd 3rd 4th 5th 6th

ALPHABET

P/F 1. A long time ago, what did people do when they wanted to write something?_____

P/F 2. Writing with pictures does not work very well. Why? _____

P/F 3. Who were the first people to use an alphabet?

_____

P/F 4. In the Semite alphabet, what did the signs stand for? _____

P/F 5. How did the Semites write the sound of "a"?_____

P/F 6. How did the Semites write the sound of "b"?_____

P/F 7. How many letters were in the Semite alphabet?_____

P/F 8. What did the Greeks and Romans do to the Semite alphabet?_____

P/F 9. What is the most common way to write a language?

_____

ANIMALS

P/F 1. About how many kinds of animals do scientists know of? _____

P/F 2. When scientists find that animals are alike, what do they do? _____

P/F 3. Scientists have divided all the animals into how many main groups? _____

P/F 4. Name an animal given in the story that does not have a backbone. _____

P/F 5. Name an animal given in the story that has a backbone. _____

P/F 6. Do fish have backbones? _____

P/F 7. How fast can the eagle fly? _____

P/F 8. Can a fish drown? _____

P/F 9. How does the walking fish get air to breath?

_____

FOLK TALE

P/F 1. What is a folk tale? _____

P/F 2. Who are the main characters in some folk tales?

_____

P/F 3. What is unusual about the animals in folk tales?

_____

P/F 4. What is the name of this folk tale? _____

_____

P/F 5. Where did the Fox get the fish? _____

P/F 6. What did the Fox tell the Bear not to worry about?

_____

P/F 7. How did the Fox tell the Bear to pull out his tail?

_____

P/F 8. Who was fooled in this story? _____

FOOD

P/F 1. What was the first country you visited?_____

P/F 2. Are there any people who eat monkeys? _____

P/F 3. What do people at the North Pole eat?_____

P/F 4. What did your friends use to kill the seal?_____

P/F 5. Why didn't your friends cook the seal?_____

_____

P/F 6. In Arabia, where is the milk kept? _____

P/F 7. Why do your friends like sour milk?_____

_____

P/F 8. What was your special dinner in China?_____

P/F 9. What color is bird's-nest soup?_____

P/F 10. Where does the swallow build its nest? _____

P/F 11. How did your friends get the nest?_____

P/F 12. What might happen to your friends when they go to get the bird's nest? _____

PLANTS

P/F 1. Can a plant eat meat?_____

P/F 2. Where does Venus's-flytrap grow? _____

P/F 3. How tall does Venus's-flytrap grow?_____

P/F 4. What part of Venus's-flytrap closes like a trap?

_____

P/F 5. What shape does a pitcher plant leaf have?

_____

P/F 6. What is inside the leaves of a pitcher plant?

_____

P/F 7. Why do insects land on the leaves of pitcher plants? _____

P/F 8. What causes an insect to fall into the water of a pitcher plant? _____

P/F 9. What happens to an insect that falls into the water of a pitcher plant? _____

P/F 10. What insect makes its home in the pitcher plant?

_____

P/F 11. Where do most plants get their food? _____

P/F 12. What makes Venus's-flytrap and the pitcher plant special? _____

_____Total P

APPENDIX E

BUILDING AND SHAKING THE EARTH

FOURTH GRADE CLASSROOM LISTENING QUESTIONS

A LETTER ACROSS TIME

SIXTH GRADE CLASSROOM LISTENING QUESTIONS

BUILDING AND SHAKING THE EARTH

On February 20, 1943, a Mexican farmer went to work in his fields. The sky was clear over the village of Paricutin. Late in the afternoon the farmer heard a low rumble. He saw a column of smoke only 6 centimeters wide rising from a small hole in the ground. Three hours later, the hole was a pit 10 meters across. Clouds of ash billowed from the pit. Red-hot stones were thrown into the air. The low rumble grew into a terrifying noise. The next day, in place of the pit stood a hill 12 meters high! A volcano was born! During the next 9 years it grew over 400 meters above the farmer's ruined fields.

Whenever you push against or pull on an object, you use force. You use force to push a tack into a wooden board. If you use a machine, such as a hammer, you make the force greater.

But forces and pressures inside the earth are much greater than any machine can make. Forces within the earth push and pull on the earth's crust.

Most paved streets, sidewalks, and playgrounds have cracks. Pressure from under the ground causes some of the cracks. For example, a tree root might grow under a sidewalk. The growing root presses against the sidewalk and can crack it. Pressure in the earth's crust also causes cracks. If the rock in the crust moves or slides along a crack, the crack is called a fault. Many faults are deep in the crust. We cannot see them. But some faults are visible on the earth's surface. Sometimes land along a fault moves from side to side instead of up and down.

A strong earthquake took place in southern Italy in 1980. A large fault near southern Italy is at the boundary between two plates. Thousands of people were killed. In less than one minute, more than 100 towns crumbled!

Scientists use instruments to help them measure earthquakes . These instruments measure the slightest change in the shape of the land. Scientist can then assign a number from 1 to 12 to the earthquake. A one is not felt by people. A 12 destroys everything made by people.

Scientists around the world are tying to learn how to predict earthquakes. They think certain things may happen before an earthquake strikes. The ground may bulge. The water level in a well may suddenly change. Small earthquakes, called tremors, could mean a strong earthquake is coming. Some scientists think animals behave strangely before an earthquake.

When thick, gray smoke rises from a house, you know something is happening. The house is on fire. When smoke or ash rises from a volcano, you know something is happening deep inside the earth. The inside of the earth is hot enough to melt rock, called magma. Pressure squeezes the magma up through the crust just as you squeeze toothpaste through to the opening. Magma that comes out onto the surface of the earth is called lava. After many eruptions, lava and rocks build a mountain called a volcano.

When a volcano erupts, it is deadly to all living things nearby. Hot, choking ask kills people and animals and destroys or damages crops. After many years, hardened lava changes into dark soil, which makes good farmland. Some countries, such as the United States, Japan, and Italy are learning to use heat from volcanoes to make electricity.

FOURTH GRADE CLASSROOM LISTENING QUESTIONS

1. What did the farmer hear?

2. What do you use when you push or pull against an object?

3. What is a crack in the earth that rock moves along called?

4. What direction can the land move when the boundary between two plates move?

5. How long did it take the earthquake in Southern Italy to destroy hundreds of towns?

6. Scientists measure earthquakes by using numbers. Describe an earthquake that rates as a one.

7. What are small earthquakes called?

8. Name one sign used to predict earthquakes?

9. What are magma and lava?

10. What good can come from a volcano?

11. Did you already know any part of this lesson?

A Letter across Time

Dear Friend across Time:

I do not know if this letter will ever be found, but I shall write it so that one day, perhaps, someone will find it and know a little about our life here in Sparta, Greece, during this year following the Spartans' defeat at Thermopylea Pass. I am telling my letter to our village scribe. He is the person who writes things for other people. For, you see, not all of us learn to write. Perhaps someday I will.

From the age of seven I have been in a big, stone building called a barracks. I am with many other boys my age. I will be in a place like this until I am thirty years old and no longer needed by my country. You see, our country for hundreds of years has been beset upon by outside invaders. Once we almost lost our land to these intruders. This alarmed one of our kings who said that from now on we must be prepared to defeat any enemy. Therefore, each male child must learn warfare from the age of seven. That includes the children of the princes and the children of the serfs. It is not a very pleasant life, but it must be done.

We are taught four main ideas. We hear them all of the time, and we must never forget them. They are to be brave, to obey, to conform or be like everyone else, and to lead a simple life. These are ideas for our souls. We are also taught that our bodies are the most important things we have. We learn to make them strong and keep them healthy.

I learn some things that others in another city state might think are wrong. For example, I have been stealing food for almost a year. It is part of my training. I learn to sneak into places without being caught. I am, in a sense, learning to go behind enemy lines, only in this case the enemy might be my neighbor. If I am caught, I am flogged. The whip hurts very much; I learned the hard way!

I have a sister at home. She will not go to war. Like the other women she has more freedom than the men. She works very hard; but if she had lived in Sparta years ago, she would have spent her time listening to poets who used to come to the village well to sing great poems. But that happens no more. Our new king does not encourage poets or any artists, for that matter. We live for one cause: protecting our nation. While my life, as well as that of my sister, has no entertainment, I do not mind, for it has to be that way. When I am 30, and if I am not killed in a war, I shall be free to farm or run a small booth at the bazaar. For now I must do what I am told. I have the bravery and the obedience to do so. I want to be like everyone else. I do not want to be different.

Your friend,
Theopolus

SIXTH GRADE CLASSROOM LISTENING QUESTIONS

1. Why did the King decide that boys should be trained in warfare?

2. How does Theopolus feel about his life.

3. What beliefs of his King seem to be part of Theo's thinking?

4. Why does Theo steal?

5. What happens if he gets caught?

6. Why are artists discouraged from practicing their arts?

7. What does the word, "conform" mean?

8. What does "flog" mean?

9. What kind of life will Theo have after his years of service?

10. What might take place in the center of Theo's village?

11. Did you already know any part of this lesson?

"A Letter across Time" and sixth grade classroom listening questions were adapted from:

Ketchum, Peter J. (1980). Hayes exercises in reading Comprehension: Grade 6 (Book F). Printed in U. S. A. Hayes School Publishing Co., Inc.

VITA

Peggy Alexander Kerr

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE DEVELOPMENT OF A SINGLE FACTOR LISTENING COMPREHENSION TEST

Major Field: Applied Behavioral Studies

Biographical:

Personal Data: Born in Rush Springs, Oklahoma, September 19, 1941, the daughter of Basil and Katie Alexander.

Education: Graduated from Edmond High School, Edmond, Oklahoma, in May 1959; received Bachelor of Arts Degree from Central State University in May 1962; received Master of Education Degree from Central State University in July 1976; completed requirements for the Doctor of Philosophy Degree at Oklahoma State University in May, 1989.

Professional Experience: Teacher of English and Mathematics in the High Schools of Burlington and Wray, Colorado, 1962 to 1966; teacher of learning disabled children in the Elementary Schools of Putnam City, Oklahoma, 1976-77, and Deer Creek of Edmond, 1977-83. Instructor and Coordinator of Psychometry and School Psychology Certificate Programs at Central State University, 1987 to present.