A comparison of two language screening instruments in two populations

Karen Kumpula-Lacey
Portland State University

Recommended Citation
Dissertations and Theses. Paper 3178.
https://doi.org/10.15760/etd.3167

Title: A Comparison of Two Language Screening Instruments in Two Populations.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

Robert Ca Steel, Chair

Mary Gordon

Steve Brannan

Screening has been suggested as the most efficient method to find students with potential language problems (Neidecker, 1980). Based on the need for a standardized adolescent language screening tool, Prather, Breecher, Stafford, and Wallace (1980) developed the Screening Test of Adolescent Language (STAL). This is a six to eight minute test with twenty-three items which examine vocabulary, auditory memory span, language processing, and proverb explanation. Following an item analysis of the STAL, Prather, Brenner, and Hughes (1981)
derived the **Mini-Screening Test of Adolescent Language** (M-STAL). This test contains five items from the STAL and requires one minute to administer.

The purpose of the present study was to determine how the results of the M-STAL correlate with the STAL in the sixth through ninth grade students compared to the tenth through twelfth grade students. Furthermore, the items of the M-STAL were correlated to the STAL in a multiple regression analysis.

The 207 sixth through ninth grade subjects and the 116 tenth through twelfth grade subjects were drawn from Clatsop County, Oregon, and tested with both instruments individually.

The results of the study showed no significant difference in the number of false positives in the two age groups. The percentage of false negatives was also similar in the groups.

In the Multiple Regression analysis the highest correlation of one of the five items with the STAL was .35 indicating, according to Guilford (1965), a low correlation, a small yet significant relationship. An Index of Determination ($R^2$) indicated a shared variance between the M-STAL items to the STAL score to be 13 percent (sixth through ninth) and 23 percent (tenth through twelfth). This implies a 77 percent (tenth through twelfth) and 87 percent (sixth through ninth) domain variability which was left unaccounted for in this study.

In addition, by examining the STAL failures in the sixth through ninth population the M-STAL missed 29 percent of the students who would have failed the STAL.

By examining the discrepancy between Prather et al. results and
the present study results, the low correlations of the M-STAL items to the STAL, the high percentage of domain variability between the tests, and the percentage of students left undetected, this writer does not support the use of the M-STAL.
A COMPARISON OF TWO LANGUAGE SCREENING INSTRUMENTS IN TWO POPULATIONS

by

KAREN KUMPULA-LACEY

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN SPEECH COMMUNICATION: with an emphasis in
SPEECH PATHOLOGY/AUDIOLOGY

Portland State University
1982
TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Karen Kumpula-Lacey presented April 29, 1982.

Robert Castéel, Chair

Mary Gordon

Steve Brannan

APPROVED:

Theodore Grove, Chair, Department of Speech Communication

Stanley E. Rauch, Dean, Graduate Studies and Research
ACKNOWLEDGEMENTS

My initial appreciation must be expressed to Elizabeth Prather for enthusiastically allowing me to use her test instrument in my study. Dr. Casteel provided massive portions of time in editing, organizing, and refining my study. For all those hours of time, plus your support, thank you. Mary Gordon, Dr. Brannan, and Ted Grove also refined my thesis by providing information, asking questions, and making suggestions. Their input was vital in pointing out areas of ambiguity. The immediate expressions of "congratulations" by the PSU staff upon completion of thesis orals also was much appreciated.

The students in the schools made the study a delight. All 323 deserve a big thank you for not only taking the test but also being so cooperative, talkative, and fun.

The Special Services directors, counselors, secretaries, librarians, and teachers were extremely helpful in setting up the testing schedule. The time they gave voluntarily, particularly in Knappa, Oregon, was exceptional.

My speech pathology buddy and close friend, Denice Milholland, also must be acknowledged. For all the time we spent "drilling for a test," laughing, crying, sharing secrets, and supporting one another, thank you.

Then, to the group of people I owe my greatest appreciation, my family—Mike Lacey, Lily, Don, Carol, Paul, and David. Beginning with my brother David, thanks for bribing, bugging, and begging your
seventh grade friends to take my test. Your effort made my sample quite large! Paul, thanks for being gracious enough to put up with a sister in high school again. You faked it well!

And to the mighty foursome—Mike, Carol, Dad, and Mom—who told me to "hang in there," as I threatened to quit at least a thousand times, you made it all worth-while. With my Mom's sense of confidence in me, my Dad's sense of balance ("have fun, too"), Carol's long-standing friendship, and Mike's listening ears and constant support, I couldn't lose.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .......................................................... iii
LIST OF TABLES ............................................................... vii

CHAPTER

I INTRODUCTION AND STATEMENT OF PURPOSE ..................... 1
   Introduction ............................................................. 1
   Statement of Purpose .................................................. 3
   Definition of Terms .................................................... 3

II REVIEW OF THE LITERATURE ........................................... 5
   Background to Language Assessment of Adolescents ............... 5
   Assessment Instruments Used with Adolescents .................. 7
   STAL and M-STAL ....................................................... 12

III METHODS ................................................................. 15
   Subjects .................................................................. 15
   Testing Environment ................................................... 15
   Procedures ............................................................... 15
   Instruments .............................................................. 17
   Data Analysis ........................................................... 18

IV RESULTS AND DISCUSSION ............................................. 20
   Results .................................................................. 20
   Discussion ............................................................... 24
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Cut-Off Scores of the STAL</td>
<td>18</td>
</tr>
<tr>
<td>II</td>
<td>Comparison of False Positives in 6th through 9th and 10th through 12th Grade Data by Chi Square Analysis</td>
<td>20</td>
</tr>
<tr>
<td>III</td>
<td>Comparison of False Negatives in 6th through 9th and 10th through 12th Grade Data by Percentage Figures</td>
<td>21</td>
</tr>
<tr>
<td>IV</td>
<td>Results of Multiple Regression Analysis Predicting Total STAL Score from the Five Items of M-STAL in 6th through 9th Grade Population (N=207)</td>
<td>22</td>
</tr>
<tr>
<td>V</td>
<td>Results of Multiple Regression Analysis Predicting Total STAL Score from the Five Items of M-STAL in 10th through 12th Grade Population (N=116)</td>
<td>23</td>
</tr>
<tr>
<td>VI</td>
<td>M-STAL and STAL Results from Prather et al. (1981) Study: 6th through 9th Grade Data (N=425)</td>
<td>25</td>
</tr>
<tr>
<td>VII</td>
<td>M-STAL and STAL Results from Present Study: 6th through 9th Grade Data (N=207)</td>
<td>25</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

Introduction

The identification of adolescents who are in need of language intervention is the first step toward remediation. A full scale language assessment requires an investment of forty-five minutes to one and one-half hours of time per student (Baker and Leland, 1967; Kirk, McCarthy, and Kirk, 1968; Wiig and Semel, 1980b). Traditionally, the only students who receive such extensive assessment are clearly language handicapped students. Neidecker (1980) has pointed out that case finding in the secondary school is usually accomplished by teacher referral, even though these teachers are less aware of the disorders than speech-language pathologists (Phillips, 1976). Therefore, an instrument which provides a screening of language abilities of students in the sixth through twelfth grades is necessary so that the speech clinician can identify the students who need more extensive assessment.

Since 1975 it has become increasingly important for the speech-language pathologist to have an instrument to fulfill requirements of U.S. Congressional Public Law 94-142, The Education for All Handicapped Children Act of 1975. This act mandates that the services of the speech-language pathologist include assessment of children with speech-language disorders using two assessment instruments. Because
one of these instruments must be standardized, this requirement has presented a problem for speech-language pathologists working with junior high and secondary school students.

According to Brenner (1979), speech-language pathologists have screened students using subjectively chosen items or by inappropriate use of tests standardized on a younger population. According to Fisher (1981) and Middleton (1981), others have used individual sub-tests from extensive test batteries such as the Detroit Test of Learning Aptitude (Baker and Leland, 1967), the Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy, and Kirk, 1968), and the Clinical Evaluation of Language Functions (Wiig and Semel, 1980a).

In consideration of the previous information, Prather, Breecher, Stafford, and Wallace (1980) developed a screening test which examines both expressive and receptive language skills. The test, Screening Test of Adolescent Language (STAL), was designed to be used with students from the sixth through twelfth grades. It examines vocabulary, auditory memory span, verbal explanation, and language processing. The test requires six to eight minutes to administer.

Clearly, this is a markedly more efficient tool than earlier diagnostic batteries, but as an all school screening device it is time consuming. Prather, Brenner, and Hughes (1981), therefore, developed a Mini-Screening Test of Adolescent Language (M-STAL), which may be administered in one minute. They derived this test following an item analysis of the STAL.

Though Prather et al. (1981) state that the M-STAL was developed for sixth through twelfth grade students, they have not compared the
results of the M-STAL and STAL with students in tenth through twelfth grades.

Statement of Purpose

The purpose of this investigation was to determine how the results of the M-STAL correlate with the STAL in screening sixth through ninth grade students as compared to tenth through twelfth grade students.

This investigation will seek to answer the following questions:

(a) Is there a significant difference between the sixth through ninth grade students and the tenth through twelfth grade students who pass the M-STAL but fail the STAL?

(2) Is there a significant difference between the sixth through ninth grade students and the tenth through twelfth grade students who fail the M-STAL but pass the STAL?

(3) What is the predictive value of the M-STAL items to the total STAL score in the sixth through ninth and the tenth through twelfth grade students?

Definition of Terms

For the purpose of this study the following operational definitions were utilized:

Language: Oral language is a code with structured properties characterized by a set of rules for producing and comprehending spoken utterances (Hixon, Shriberg, and Saxon, 1980).

Screening: Screening is a systematic procedure to identify provisionally those students from a population who manifest, or are likely to manifest, an attribute which is judged to require special attention; it is used to identify those students needing a more complete assessment (Hill, 1970).

Children with special learning disabilities exhibit a disorder in one or more of the basic psychological processes involved in understanding or using spoken or written language. These may be manifested in disorders of listening, thinking, talking, reading, writing, spelling, or arithmetic. They include conditions which have been referred to as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, developmental aphasia, etc. They do not include learning problems which are due primarily to visual, hearing, or motor handicaps, to mental retardation, emotional disturbances, or to environmental disadvantage.

Adolescence: The period of life beginning in the sixth grade and terminating upon completion of the twelfth grade.
CHAPTER II

REVIEW OF THE LITERATURE

While the language of young children has been studied and assessed extensively over the years, the language abilities of the adolescent have remained relatively unexamined (Hamill, 1980). According to Byrne and Shervanian (1978), there are students at the secondary level who have language problems needing intervention. In addressing this issue, Neidecker (1980) has suggested screening as the most efficient method to finding students with potential language problems.

The present review of the literature, relative to the screening of language problems in adolescents, will be presented under three subtopical headings: Background to Language Assessment of Adolescents; Assessment Instruments Used with Adolescents; and the STAL and M-STAL.

Background to Language Assessment of Adolescents

Historically, evaluation of the language of the adolescent has been minimal due to lack of information on language development beyond the elementary level (Thorum, 1981). Due to this lack of information, reliable, valid, and educationally relevant measures did not exist (Rosenberg, 1970). Consequently, when an adolescent was evaluated in the area of language, a vocabulary, aphasia test, or children's
language test was used (Sitko and Gillespie, 1975).

The tests used were such tests as the Peabody Picture Vocabulary Test (Dunn, 1956), Language Modalities Test for Aphasia (Wepman and Jones, 1961), and the Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy, and Kirk, 1968). Until the midseventies when language skills in the adolescent would be more extensively examined, these tests comprised the adolescent assessment (Thorum, 1981).

With creation of U.S. Congressional Public Law 91-230, The Children with Specific Learning Disabilities Act of 1969, language problems became linked with learning disabilities in elementary and secondary students (Baren, Liebl, and Smith, 1978; Gerber, 1981). According to Sitko and Gillespie (1975), research began to examine language problems not only as a symptom of learning disabilities but also as a cause. An interest in the younger learning disabled population by parents and organizations led to the development of assessment tools for those age levels (Mercer, 1979). Language intervention began to be suggested for the learning disabled child with the major emphasis still focused on the preschool or elementary levels (Drake and Cavanaugh, 1970; Bannatyne, 1971; Lyness, 1973).

Although Public Law 91-230 included the secondary student within the learning disabled definition, the development of language skills through intervention with the secondary student was minimal (Mercer, 1979). Contrary to the preschool and elementary students, the language problems of the secondary student were not considered developmental. The intervention strategies, therefore, were aimed at teaching them how to compensate for their language deficiencies (Wepman, Cruickshank,

With creation of Public Law 94–142, *The Education for All Handicapped Children Act of 1975*, the identification of children with speech and language problems was mandated. Following this act, research in the area of adolescent language began to emerge. Menyuk (1977) stated that the advent of adolescence did not terminate one's ability to acquire additional knowledge in language. Curtiss, Krashner, Fromkin, Rigler, and Rigler (1973) and Myerson (1975) documented continued language development beyond puberty until the age of seventeen years. Loban's (1976) longitudinal study of 211 students from kindergarten to twelfth grade found language development to extend through twelfth grade. Thus, research began to stress the importance of language intervention in the adolescent population (Loban, 1976; Menyuk, 1977).

In order to identify the secondary student with language problems the speech pathologist needed standardized instruments. Due to the void in the area of both screening and extensive language tests for adolescents, speech-language pathologists used a method involving interpolation of vocabulary, aphasia, achievement, and children language tests (Kayser, 1981).

**Assessment Instruments Used with Adolescents**

The instruments used earlier, *Peabody Picture Vocabulary Test* (PPVT), *Language Modalities Test for Aphasia* (LMTA), and the *Illinois Test of Psycholinguistic Abilities* (ITPA) continued to be used until the midseventies, along with a variety of other tests. Some of the
other tests included: **Detroit Test of Learning Aptitude** (Baker and Leland, 1967), **Peabody Individual Achievement Test** (Dunn and Markwardt, 1970), and the **Utah Test of Language Development** (Mecham, Jex, and Jones, 1967). A brief description of these tests will follow in order to give the reader background on these tools.

The PPVT (Dunn, 1956) is an instrument designed to measure receptive vocabulary capabilities in children two and a half to eighteen years. The test consists of 150 plates with four pictures in which one corresponds to a word orally presented by the examiner. The raw score may be converted into three types of derived scores: mental age, intelligence quotient, and percentile ranking. The test was standardized on 4,012 subjects over the age range of two years through eighteen years. The administration time involves fifteen minutes per student.

The LMTA (Wepman and Jones, 1961) is an instrument designed to evaluate aphasic adults. The major sensory modalities of vision and audition and the major pathways for speech and writing are examined. The test requires one to three hours to administer. No normative data are provided.

The ITPA (Kirk et al., 1968) is an instrument composed of twelve subtests to be used with children two to ten years old. The subtests are designed to fit a communication model composed of three constructs: (1) channels of communication (visual and auditory), (2) psycholinguistic processes (receptive, expressive, or organizing), and (3) levels of organization (representational and automatic). The test may be administered in forty-five minutes to one hour. It yields a scaled
or an age score for each subtest and an overall psycholinguistic age can be obtained from the raw score.

The Detroit Test of Learning Aptitude (DTLA) is a test designed to assist in the assessment of abilities in reasoning and comprehension, practical judgment, verbal ability, time and space relationships, number ability, auditory attention span, visual attention ability, and motor development. It yields a mental age and intelligence quotient for the age range of three through nineteen. The test was standardized on 150 students of the Detroit Public Schools at each age level (three-month intervals from three to nineteen years). The test may be administered in one hour.

The Peabody Individual Achievement Test (PIAT) is designed to provide a wide-range measurement of achievement in the areas of mathematics, reading, spelling, and general information. The test has normative data for the age range four years to adult. It was standardized on 2,880 subjects from kindergarten to grade twelve. The test may be administered in forty minutes to one hour.

The Utah Test of Language Development (UTLD) is a fifty-one item test which contains approximately five items for each age level from one year to fifteen years. The items are presented in developmental sequence. The total score yields a language-age-equivalent which can be compared to the child's chronological age. The test was standardized on 273 children in Utah. The test may be administered in fifteen minutes.

As these tests were used with adolescents, speech-language pathologists began to criticize the lack of normative data on adoles-
cents, poor validity and reliability measures, and weak theoretical
development of these instruments (Irwin, Moore, and Ramps, 1972;
McCarthy, 1975; Sitko and Gillespie, 1975; Hallaton and Kauffman,
1976). Speech-language pathologists began to request language assess­
ment tools made for adolescents which were based on current language
theory (Sitko and Gillespie, 1975).

One of the first tests to appear was the **Wiig-Semel Test of
Linguistic Concepts** (Wiig and Semel, 1974). This test was designed to
evaluate the comprehension of fifty linguistic concepts in children
six through fourteen years. The test examined: (1) comparatives,
(2) passive relationships, (3) temporal relationships, (4) spatial
relationships, and (5) familial relationships. The test was standard­
ized on 240 grade school children, thirty from the first through
eighth grade. It requires ten to fifteen minutes to administer.

Another test that emerged was the **Test of Syntactic Abilities**
(Quigley, 1978). This test was designed to provide diagnostic meas­
ures on deaf students in written comprehension and use of syntactic
structures of Standard English in ten- through eighteen-year-old
children. It consists of a battery of twenty individual diagnostic
tests, each containing seventy multiple choice written items. Each
subtest requires twenty minutes to administer. The test was standard­
ized on 450 profoundly deaf students.

In 1979, the **Fullerton Language Assessment Test for Adolescents**
by Thorum became available to speech-language pathologists. It was
designed to provide information concerning the language processing
abilities of students, eleven to eighteen years. There are eight sub-
tests which examine auditory synthesis, morphology, oral commands, convergent production, divergent production, syllabication, grammatic competency, and idioms. A numerical scoring system is used to determine language proficiency, while a percentage score provides frustration level, instruction level, and competence level designations. The test was standardized on seventy-five adolescents for each age level from age eleven to eighteen. The test requires thirty to forty-five minutes to administer.

In 1980, Wiig and Semel developed the Clinical Evaluation of Language Functions (1980a). It contains two elements: a screening test and a diagnostic battery. The screening test has two levels, the elementary level (kindergarten to fifth grade) and an advanced level (fifth grade through twelfth grade). This screening test provides a measure of language processing and production abilities. The size of the standardization sample was 1,405 children. The screening test requires fifteen minutes to administer. The battery consists of thirteen subtests divided into three sections: processing subtests, production subtests, and supplementary subtests. It requires one and one-half hours to administer the entire battery. The test was standardized on 159 children ranging in grade level from kindergarten through grade twelve.

Also in 1980, Hamill made available his test, Test of Adolescent Language (TOAL). It was designed to provide information on language abilities of students in the seventh through twelfth grades. It consists of eight subtests in which four assess vocabulary abilities and four assess grammar abilities. A general index of overall language
abilities plus composite scores for the following ten areas is provided: listening, speaking, reading, writing, spoken language, written language, vocabulary, grammar, receptive language, and expressive language. The test was standardized on 2,733 students and required one and one-half to two hours to administer.

STAL and M-STAL

Although the instruments mentioned have been utilized to examine adolescent language problems, none of these instruments are efficient in large scale screening (Kayser, 1981; Middleton, 1981; Fisher, 1981). Prather, Breecher, Stafford, and Wallace (1980), therefore, created the Screening Test of Adolescent Language (STAL), which allowed screening of forty-five students in one day.

The STAL includes twenty-three items divided among four subtests as follows: twelve items on vocabulary, three items on auditory memory span, five items on language processing, and three items on proverb explanation. According to the authors these four subtests were selected to examine several language skills associated with language disabilities by Wiig and Semel (1976). The vocabulary subtest requires word finding and retrieval competencies. The auditory memory span subtest examines the aspect of memory involved in semantic and syntactic stimuli. The language processing subtest examines the student's ability to decode a message and use language for reasoning and problem solving. Finally, the proverb explanation subtest investigates paraphrasing and cognitive skills needed for verbal clarity.

The sample of students in the normative data of the STAL
included 206 sixth graders and 219 ninth graders in the Phoenix School District. The criteria for selection specified that the subjects could not be enrolled in special education or special reading nor have received any speech or language intervention in the past.

The authors Prather et al. (1980) selected a cut-off score for the total test which failed approximately 9 percent of those tested. Furthermore, students who passed the test but failed one subtest were also among the students they considered needed further evaluation. The preceding criteria, therefore, placed 11 to 12 percent of the normative subjects in the category needing further evaluation.

Establishment of the validity of the STAL involved thirty-eight students in the ninth grade who scored across the range from low to high on the STAL. These thirty-eight students received four subtests of the Detroit Test of Learning Aptitude (Baker and Leland, 1967): verbal opposites, auditory attention span for related syllables, verbal absurdities, and likenesses and differences. According to the authors Prather et al. (1980), the DTLA was selected as the instrument for validity correlation because it was standardized at the adolescent levels and included a large number of items. The Pearson product-moment correlation between the total STAL score and the total raw score across the four DTLA subtests was $r_{86}$.

The reliability of the test was determined by testing thirty students one month later. The Pearson product-moment correlation coefficient between these two sets of scores was $r_{98}$.

Although the administration time of the STAL is only six to eight minutes per student, an even shorter test was desirable for
large scale screening. Prather, Brenner, and Hughes (1981) consequently performed an item analysis on the STAL and derived the **Mini-Screening Test of Adolescent Language** (M-STAL). This test has two items from the vocabulary subtest, and one item from each of the other subtests on the STAL. It requires only one minute to administer and allows screening of 300 students in one day.

The population used for the normative data included 206 sixth graders, 219 ninth graders, and 122 learning disabled students in the Phoenix School District. At the sixth grade level the M-STAL had a prediction accuracy of 84 percent with 75 percent passing both tests (STAL and M-STAL) and 9 percent failing both tests. At the ninth grade level the M-STAL had a prediction accuracy of 84 percent, with 74 percent passing both tests and 10 percent failing both tests. The junior high learning disabled had a 78 percent prediction accuracy, with 36 percent passing both tests and 42 percent failing both tests.

According to the results of the M-STAL on the sixth and ninth grade regular classrooms, the speech-language pathologist may predict that 75 percent will pass the M-STAL and 25 percent fail (Prather et al., 1981).

It has been noted in the preceding review of the literature that there is a definite need for screening tests which predict potential language problems in the secondary student. Although the M-STAL allows for screening 300 students in one day, it has been tested only with a limited number of sixth and ninth graders.
CHAPTER III

METHODS

The following is a discussion of the methods utilized in the present study. The selection of the subjects, the environment, procedures, and instruments will be presented.

Subjects

The subjects in this study were drawn from Clatsop County School District and Astoria School District (State of Oregon). From this pool of 1,000 students, in the sixth through twelfth grades, the 323 students who returned parent permission slips for testing were chosen as subjects.

No attempt was made to control for intelligence, race, socio-economic level, or sex.

Testing Environment

Students were tested individually in an area outside the classroom. Although separate arrangements were made in each district, all testing environments consisted of an area with two chairs and a table.

Procedures

Parent permission slips were given to students in the two districts (see Appendix A). When the slips were returned, the subjects
were scheduled for testing.

Prior to the administration of the tests, the examiner noted the subject's name and grade. She then assigned the subject a test number on the response form to assure confidentiality. During the testing the student sat across a table from the examiner. The tape recorder, tapes, and response forms were placed on the table. Response forms, along with the lists of subjects, were placed out of the student's view.

At the initiation of the testing, the examiner explained that the screening test was not a graded school test. She also stated that some items would be repeated on the second test. The tape recorder was explained as a device used to check the examiner's scoring of the test.

In administration of the tests, the order of the presentation of the STAL and M-STAL was randomized. The standardized instructions were given as directed in the manual (see Appendix B). The examiner recorded the results according to the test directions. The vocabulary items were recorded verbatim. Those answers which were considered correct according to the test manual received a score of one while those considered incorrect received a score of zero. The auditory memory items were examined according to the items omitted, substituted, or added to the original sentence. The answer was given a score of one if it had zero to two errors. A score of zero was given if it had three or more errors. In the language processing subtest a response which included a description of "what does not make sense" and "why it does not make sense" received a score of one. Those
responses which were incomplete, incorrect, or irrelevant according to
the test manual, received a score of zero. On the final subtest, pro-
verb explanation, a response which indicated an understanding of the
proverb received a score of one, while those responses which are con-
sidered incomplete, incorrect, or irrelevant according to the test
manual, received a score of zero.

The administration of the M-STAL and the STAL was completed in
one session. The average session was seven minutes.

**Instruments**

The instruments utilized in the present study were the STAL and
M-STAL (see Appendix C). Both of these tests were designed to screen
students in the sixth through twelfth grades in the areas of receptive
and expressive language. Both tap the areas of vocabulary, auditory
memory span, language processing, and proverb explanation.

The STAL, which has twenty-three items, uses the cut-off scores
in Table I to designate the minimum passing scores for the test.
Those students who fail either the total test or any of the subtests
should be further evaluated for possible language problems (Prather et
al., 1980).

Using this table it becomes evident that a student will not pass
the test if she passes the subtests at a minimum passing score. For
example, if a sixth grader scored 5 on the vocabulary subtest, 1 on
the auditory memory span, 2 on the language processing, and 1 on
proverb explanation, she would receive a score of 9, indicating a test
failure. Furthermore, if a ninth grader scored a 13 total test score,
she would fail if she had a vocabulary score that was 5 rather than 6. Consequently, a student must not only pass all subtests with a minimum passing score but also must pass two more items (if sixth through eighth grader) or three more items (if ninth through twelfth grader), in order to pass the total test.

On the M-STAL, which is a five item test, those students who fail one or more of the items are classified as needing further evaluation (Prather et al., 1981).

**Data Analysis**

The results of the STAL and the M-STAL were analyzed with the Chi Square analysis to determine if there were significantly more false positives in either the sixth through ninth or tenth through twelfth grade students. Furthermore, a Multiple Regression analysis
was done to determine the predictive value of the five items of the M-STAL to the total score of the STAL in the two age groups separately.
CHAPTER IV

RESULTS AND DISCUSSION

Results

The Chi Square ($X^2$) analysis was used in order to answer the question: Is there a significant difference between the sixth through ninth grade students and the tenth through twelfth grade students who fail the M-STAL but pass the STAL? The M-STAL and STAL results were dichotomized into two categories: "fail M-STAL/pass STAL" and "all other combinations" (see Table II). The $X^2$ value computed from these data was .8011, indicating that the $X^2$ value did not reach significance at the .05 level.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Fail M-STAL/</th>
<th>All other combinations</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pass STAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6th-9th</td>
<td>20 (22.4)</td>
<td>187 (184.6)</td>
<td>207</td>
</tr>
<tr>
<td>10th-12th</td>
<td>15 (12.6)</td>
<td>101 (103.4)</td>
<td>116</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>288</td>
<td>323</td>
</tr>
</tbody>
</table>

$x^2 = .8011$, df=1

With 1 degree of freedom, an $X^2$ value of 3.84 is required for significance at the .05 level.
The false negatives were computed in order to answer the question: Is there a significant difference between the sixth through ninth grade students and the tenth through twelfth grade students who pass the M-STAL but fail the STAL? In this comparison the M-STAL and STAL results were dichotomized into two categories: "pass M-STAL/fail STAL" and "all other combinations" (see Table III). Although the small number of false negatives (four for sixth through ninth and two for tenth through twelfth) eliminated the use of the Chi Square analysis, the percentages of the false negatives in both groups indicated similar performances (1.93 percent for the sixth through ninth grades and 1.72 percent for the tenth through twelfth grades).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pass M-STAL/fail STAL</th>
<th>All other combinations</th>
<th>Total</th>
<th>False Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th-9th</td>
<td>4</td>
<td>203</td>
<td>207</td>
<td>1.93%</td>
</tr>
<tr>
<td>10th-12th</td>
<td>2</td>
<td>114</td>
<td>116</td>
<td>1.72%</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>317</td>
<td>323</td>
<td>1.86%</td>
</tr>
</tbody>
</table>

Finally, a Multiple Regression analysis was performed on the five items of the M-STAL to the total STAL score in both age groups in order to answer the question: What is the predictive value of the M-STAL items to the total STAL score in the sixth through ninth and
the tenth through twelfth grade students? This analysis mathematically assigned a weight to each item and indicated the contribution of each predictor variable (items 1-5) in predicting STAL scores. Using the sixth through ninth grade data with this analysis, item 5 (proverb) demonstrated the highest correlation value of .271. Item 2 (vocabulary) had a correlation coefficient of .265. Items 3 (auditory memory span) and 4 (language processing) illustrated negligible correlations. Item 1 made no distinctions at all, as it was correctly answered by all the students in this sample (see Table IV).

TABLE IV

RESULTS OF MULTIPLE REGRESSION ANALYSIS PREDICTING TOTAL STAL SCORE FROM THE FIVE ITEMS OF M-STAL IN 6TH THROUGH 9TH GRADE POPULATION (N=207)

<table>
<thead>
<tr>
<th>Variables</th>
<th>R</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1 (vocabulary)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Item 5 (proverb explanation)</td>
<td>.271</td>
<td>.073</td>
</tr>
<tr>
<td>Item 2 (vocabulary)</td>
<td>.265</td>
<td>.127</td>
</tr>
<tr>
<td>Item 3 (auditory memory span)</td>
<td>.026</td>
<td>.128</td>
</tr>
<tr>
<td>Item 4 (language processing)</td>
<td>.002</td>
<td>.129</td>
</tr>
</tbody>
</table>

*Item 1 made no distinctions, as all students accurately answered the item.

A second part of the Multiple Regression analysis (Table IV) involves an Index of Determination (R²) indicating the amount of shared variance between the M-STAL items and the total STAL score.
The cumulative $R^2$ for this sample is equal to .129, indicating that more than 87 percent of the variance between the M-STAL items and the total STAL score was not accountable in this sample.

Using the tenth through twelfth grade data in a Multiple Regression analysis, item 3 had the highest correlation value at .347. Item 5 demonstrated a correlation at .291. Items 2 and 4 had negligible correlation values, while item 1 could not be computed (see Table V). Item 1 made no distinctions based on the accurate performance by all the students.

**TABLE V**

RESULTS OF MULTIPLE REGRESSION ANALYSIS PREDICTING TOTAL STAL SCORE FROM THE FIVE ITEMS OF M-STAL IN 10TH THROUGH 12TH GRADE POPULATION (N=116)

<table>
<thead>
<tr>
<th>Variables</th>
<th>R</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1 (vocabulary)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Item 3 (auditory memory span)</td>
<td>.347</td>
<td>.120</td>
</tr>
<tr>
<td>Item 5 (proverb explanation)</td>
<td>.291</td>
<td>.229</td>
</tr>
<tr>
<td>Item 4 (language processing)</td>
<td>.104</td>
<td>.229</td>
</tr>
<tr>
<td>Item 2 (vocabulary)</td>
<td>.006</td>
<td>.230</td>
</tr>
</tbody>
</table>

*Item 1 made no distinctions, as all students accurately answered the item.

The Index of Determination ($R^2$) for this sample was .230 (cumulative is equal to .230; see Table V). This indicates that 77 percent of the variance between the M-STAL items and the total STAL score was
Discussion

Although the present study results comparing the sixth through ninth grade students and the tenth through twelfth grade students show essentially the same performance on the M-STAL and the STAL, a difference is noted by comparing the present study to Prather et al. outcomes. Results previously obtained from the M-STAL with 206 sixth grade students and 219 ninth grade students (Prather et al., 1981) were compared with the results from the present study with sixth through ninth grade students. Tables VI and VII numerically describe the following data: number of false negatives, those adolescents who passed the M-STAL but failed the STAL; number of false positives, those adolescents who failed the M-STAL but passed the STAL; and those scores in which both test results are in agreement (pass M-STAL/pass STAL, fail M-STAL/fail STAL). Clearly, the most critical finding when comparing the two studies is that the false negatives were about the same, while the number of fail/fails was one-half as many in the current study as compared to Prather et al. study. This would indicate that the fewer the students who are detected, the greater the proportion who would remain undetected (pass M-STAL/fail STAL). In this study 29 percent of those who needed to be further tested went undetected by the M-STAL.

Because the six to eight minute STAL is a screening tool itself, the practicing clinician must decide if the time gained through the shorter version, M-STAL, compensates for the loss of 29 percent of the
### TABLE VI

**M-STAL AND STAL RESULTS FROM PRATHER ET AL. (1981)**

*STUDY: 6TH THROUGH 9TH GRADE DATA (N=425)*

<table>
<thead>
<tr>
<th></th>
<th><strong>STAL performance</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Pass</strong></td>
<td><strong>Fail</strong></td>
</tr>
<tr>
<td><strong>Pass</strong></td>
<td>6th (75%)</td>
<td>6th (3%)</td>
</tr>
<tr>
<td><strong>M-STAL performance</strong></td>
<td>9th (74%)</td>
<td>9th (2%)</td>
</tr>
<tr>
<td><strong>Fail</strong></td>
<td>6th (13%)</td>
<td>6th (9%)</td>
</tr>
<tr>
<td></td>
<td>9th (13%)</td>
<td>9th (10%)</td>
</tr>
</tbody>
</table>

### TABLE VII

**M-STAL AND STAL RESULTS FROM PRESENT STUDY:**

*6TH THROUGH 9TH GRADE DATA (N=207)*

<table>
<thead>
<tr>
<th></th>
<th><strong>STAL performance</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Pass</strong></td>
<td><strong>Fail</strong></td>
</tr>
<tr>
<td><strong>Pass</strong></td>
<td>6th-9th (83%)</td>
<td>6th-9th (2%)</td>
</tr>
<tr>
<td><strong>M-STAL performance</strong></td>
<td>6th-9th (10%)</td>
<td>6th-9th (5%)</td>
</tr>
<tr>
<td><strong>Fail</strong></td>
<td>6th-9th (10%)</td>
<td>6th-9th (5%)</td>
</tr>
</tbody>
</table>
students who would have failed the longer STAL.

The disparity of results between Prather's Arizona study (1981) and the present study may be due to many factors. One source of variance is the difference in racial distribution of the two samples. Prather's sixth grade sample (N=206) consisted of elementary schools in the Phoenix area. Her racial distribution was as follows: 67 percent white, 21 percent Hispanic, 9 percent black, 2 percent native American, and 1 percent Oriental. In Prather's ninth grade population students were selected from eleven campuses in the Phoenix Union High School District. This group included 219 students within the following racial distributions: 55 percent white, 26 percent Hispanic, 15 percent black, 2 percent native American, and 2 percent Oriental.

In the present study the sample consisted of students from three high schools and four grade schools in Clatsop County, Oregon. The racial distribution, according to the categories stated by Prather, was the following: 99 percent white and 1 percent Oriental in the elementary group; and 98 percent white and 2 percent Oriental in the high school group.

A second source of variance may be the scoring of the results. Due to the wide range of responses from students not addressed in the manual this researcher chose to use Roget's Thesaurus (Lewis, 1970) as a supplementary guide for the vocabulary items. Further, the question of whether the STAL guide provides only examples of a few correct answers or the "only correct answers" is not clearly defined. This same problem in determining correctness of response is even more crucial in the language processing subtest in which a student is
expected to provide a "what" and "why" response. All of the examples of correct responses in the manual contain complete sentences with Standard English grammar. Approximately 40 percent of the students in this sample used incomplete or grammatically incorrect sentences, yet correctly stated the "what" and "why" elements. Since the purpose of this subtest was to screen a student's ability to process and verbally express "what does not make sense and why it does not make sense," this investigator examined responses for only the "what" and "why" elements. Most definitely the interpretation of the manual by clinicians could cause a wide discrepancy in judgments.

It will be recalled that the third reason for this study was to examine the predictive value of the M-STAL (one minute screening tool) to the STAL (six to eight minute screening tool). Through use of the Multiple Regression analysis, in which the five items of the M-STAL were compared to the total STAL score, no item showed a higher correlation than .35. According to Guilford (1965), correlations from .20 to .40 indicate low correlations, yet small significant relationships. Furthermore, through this same statistic an Index of Determination ($R^2$) was computed which illustrated the small overlap between the M-STAL items to the total STAL score (.129 in the sixth through ninth, and .230 in the tenth through twelfth). This $R^2$ implies that 13 percent (23 percent in the tenth through twelfth) of the M-STAL items shared variance with the STAL leaving a full 87 percent (77 percent in the tenth through twelfth) of the domain variability unaccounted for in this study. This large domain variability between the M-STAL items to the total STAL tends to cast doubt on the value of the M-STAL as a
Due to the discrepancy between Prather et al. results and the present study results, the low correlations of the M-STAL items to the STAL, the high percentage of domain variability between the tests, and the percentage of students left undetected, this writer does not support the use of the M-STAL.
CHAPTER V

SUMMARY AND IMPLICATIONS

Summary

Screening has been suggested as the most efficient method to find students with potential language problems (Neidecker, 1980). Based on the need for a standardized adolescent language screening tool, Prather, Breecher, Stafford, and Wallace (1980) developed the Screening Test of Adolescent Language (STAL). This is a six to eight minute test with twenty-three items which examines vocabulary, auditory memory span, language processing, and proverb explanation. Following an item analysis of the STAL, Prather, Brenner, and Hughes (1981) derived the Mini-Screening Test of Adolescent Language (M-STAL). This test contains five items from the STAL and requires one minute to administer.

The purpose of the present study was to determine how the results of the M-STAL correlate with the STAL in the sixth through ninth grade students compared to the tenth through twelfth grade students. Furthermore, the items of the M-STAL were correlated to the STAL in a Multiple Regression analysis.

The 207 sixth through ninth grade subjects and the 116 tenth through twelfth grade subjects were drawn from Clatsop County, Oregon, and tested with both instruments individually.

The results of the study showed no significant difference in the
number of false positives in the two age groups. The percentage of
false negatives was also similar in the groups.

In the Multiple Regression analysis the highest correlation of
one of the five items with the STAL was .35 indicating, according to
Guilford (1965), a low correlation, a small yet significant relation­
ship. An Index of Determination (R^2) indicated a shared variance
between the M-STAL items to the STAL score to be 13 percent (sixth
through ninth) and 23 percent (tenth through twelfth). This implies a
77 percent (tenth through twelfth) and 87 percent (sixth through ninth)
domain variability which was left unaccounted for in this study.

In addition, by examining the STAL failures in the sixth through
ninth population the M-STAL missed 29 percent of the students who
would have failed the STAL.

By examining the discrepancy between Prather et al. results and
the present study results, the low correlations of the M-STAL items to
the STAL, the high percentage of domain variability between the tests,
and the percentage of students left undetected, this writer does not
support the use of the M-STAL.

Implications

In the present study it appears that the practical application
of the M-STAL in predicting STAL performance should be highly guarded.
An analysis and modification of the M-STAL are most highly suggested.
Furthermore, an even better investment of time would include a more
extensive analysis of the STAL by correlating it to the Clinical
Evaluation of Language Functions (Wiig and Semel, 1980a) and Fullerton
Language Assessment Test for Adolescents (Thorum, 1979).

An item analysis of the STAL is another suggested study to determine appropriate cut-off scores for both the junior high and high school groups.

In addition, it is of vital importance that the criteria for correctness and incorrectness of responses be clearly defined for minimal discrepancy of interpretation. Consequently, it is recommended in future revision of the manual that clearer instructions be supplied as to the range of acceptable responses, as well as delineating the types of responses which are not acceptable.
BIBLIOGRAPHY


KAYSER, V., Discussion about tests used with adolescents in the Portland area, Krestin District Annex (1981).


PRATHER, E., Correspondence through letter, Phoenix, Arizona (1981).


September 15, 1981

Dear Parent or Guardian:

I am a Portland State University graduate student doing a research project in the Speech and Hearing Sciences. The purpose of the study is to find a simple way to determine which adolescents in the school might profit from special help. No names will be used in the written results of the study.

I am requesting your permission and your son/daughter's permission for his/her involvement in my study. The project involves administering an 8 minute test to adolescents. Your son/daughter need only answer a few questions related to vocabulary, memory, and comprehension of sentences.

Please return the following slip tomorrow. Thank-you for your cooperation in my study.

Sincerely,

Karen Kumpula

I. __________________________________________________________ hereby permit my son/daughter to participate in this language study.

I. __________________________________________________________ agree to be in this language study.
APPENDIX B

ADMINISTRATION INSTRUCTIONS

Subtest I: VOCABULARY

I will say a word and then I will use it in a sentence. Give me another word that means the same thing and will fit into the sentence. Let's try an example. If a multiple word answer is given, prompt with Give me a one word answer that fits the sentence. Repeat the sentence.

Carve. Carve the turkey. Give me a word for carve.

If response is correct: Good, you used (slice, cut) for carve. Let's go on. Proceed with test items.

If response is incorrect: No, the word needed is either 'cut' or 'slice.' Slice the turkey or cut the turkey means the same as 'carve the turkey.' Let's try another one. Plate. We eat from a plate. Give me a word for plate. If response is then correct, say Good, let's go on. Proceed with the test items. The sentences may be repeated. If response is again incorrect, mark all items on this subtest as incorrect.

Subtest II: AUDITORY MEMORY SPAN

I am going to say a sentence and I want you to say it exactly the way I do. I will say it only once. Listen carefully. As each of the three sentences are read, pause at the slash marks.

Subtest III: LANGUAGE PROCESSING

I will read you a sentence. I want you to listen and tell me two things: what does not make sense and why it does not make sense. Listen carefully: I can read this only once. If the response to any sentence is incomplete, prompt with What does not make sense? or Why doesn't that make sense?

Subtest IV: PROVERB EXPLANATION

Now we are going to do the last part. I'm going to say a sentence and I want you to tell me what it means. The examiner may ask for elaboration by saying, Tell me more. The sentences may be repeated.
# SCREENING TEST OF ADOLESCENT LANGUAGE (TEST FORM)

**Five Items of M-STAL**

## SCREENING TEST OF ADOLESCENT LANGUAGE

### I. VOCABULARY

<table>
<thead>
<tr>
<th>Item</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Gigantic</td>
</tr>
<tr>
<td>2.</td>
<td>Kettle</td>
</tr>
<tr>
<td>3.</td>
<td>Unmarried</td>
</tr>
<tr>
<td>4.</td>
<td>Penalty</td>
</tr>
<tr>
<td>5.</td>
<td>Duplicate</td>
</tr>
<tr>
<td>6.</td>
<td>Bright</td>
</tr>
<tr>
<td>7.</td>
<td>Plump</td>
</tr>
<tr>
<td>8.</td>
<td>Address</td>
</tr>
<tr>
<td>10.</td>
<td>Govern</td>
</tr>
<tr>
<td>11.</td>
<td>Annoy</td>
</tr>
<tr>
<td>12.</td>
<td>Peaceful</td>
</tr>
</tbody>
</table>

**TOTAL I**

### II. AUDITORY MEMORY SPAN

<table>
<thead>
<tr>
<th>Item</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The fire drill that we had last week turned out to be the real thing.</td>
</tr>
<tr>
<td>2.</td>
<td>The school on the west side of town has more new students than our own school.</td>
</tr>
<tr>
<td>3.</td>
<td>Last night I went to a movie with my friend at the theater that takes coupons.</td>
</tr>
</tbody>
</table>

**TOTAL II**

### III. LANGUAGE PROCESSING

<table>
<thead>
<tr>
<th>Item</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The sun was shining so brightly last week on Tuesday that I had to wear my sunglasses in the movie theater.</td>
</tr>
<tr>
<td>2.</td>
<td>I went with my sister to the shoe store to buy a pair of combat boots to wear to the Junior Prom.</td>
</tr>
<tr>
<td>3.</td>
<td>After climbing up ten flights of stairs two steps at a time yesterday morning, the man finally reached the basement.</td>
</tr>
<tr>
<td>4.</td>
<td>The most recent set of identical twins born at the hospital was a girl and a boy.</td>
</tr>
<tr>
<td>5.</td>
<td>Last night after we adjusted the antenna, unplugged the television set, and changed the channel, the picture became much clearer.</td>
</tr>
</tbody>
</table>

**TOTAL III**

### IV. PROVERB EXPLANATION

<table>
<thead>
<tr>
<th>Item</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Practice makes perfect.</td>
</tr>
<tr>
<td>2.</td>
<td>Actions speak louder than words.</td>
</tr>
</tbody>
</table>

**TOTAL IV**

**TOTAL TEST SCORE**