

1991

A retrospective study of the Clinical evaluation of language functions elementary screening test (CELF-S)

Tamara Lynn Caldwell
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Speech and Hearing Science Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Caldwell, Tamara Lynn, "A retrospective study of the Clinical evaluation of language functions elementary screening test (CELF-S)" (1991). *Dissertations and Theses*. Paper 4169.
<https://doi.org/10.15760/etd.6052>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

AN ABSTRACT OF THE THESIS OF Tamara Lynn Caldwell for the Master of Science in Speech Communication: Speech and Hearing Science presented November 5, 1991.

Title: A Retrospective Study of the Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S)

APPROVED BY THE MEMBERS OF THE THESIS COMMITTEE:



Joan McMahon, Chair



Mary T. Withers



Steve Brannan

One of the more widely-used methods for pinpointing children in need of more in-depth language evaluation is screening. One language screening instrument designed to accomplish this in an effective and efficient way was the Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S) (Semel & Wiig, 1980).

The purpose of this study was to determine the ef-

fectiveness of the CELF-S in identifying those children in a second grade setting, who were in need of more thorough evaluation. This study sought to answer the following questions: 1) What is the percentage of false negatives produced by the CELF-S?, and 2) What is the percentage of false positives produced by the CELF-S?

Fifty second graders from a public elementary school were selected as subjects. Each subject passed a puretone audiometric screening and had received parental permission to be in the study. The subjects' sex, socioeconomic status, and intelligence were not controlled since this investigation sought to identify all children with possible language problems, regardless of these other factors.

Of the 50 children screened, the three students who had previously been identified as language impaired by the school's speech-language pathologist (SLP) were also identified by the CELF-S. The percentage of false negatives, therefore, was 0%. Only one student out of the 50 screened (who had not been previously tested or identified as language impaired) failed the screening. The percentage of false positives produced by the CELF-S, therefore, was 2%.

A RETROSPECTIVE STUDY OF THE CLINICAL EVALUATION
OF LANGUAGE FUNCTIONS ELEMENTARY
SCREENING TEST (CELF-S)

by

TAMARA LYNN CALDWELL

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

MASTER OF SCIENCE
in
SPEECH COMMUNICATION:
SPEECH AND HEARING SCIENCE

Portland State University
1991

TO THE OFFICE OF GRADUATE STUDIES:

The members of the Committee approve the thesis of
Tamara Lynn Caldwell presented November 5, 1991.

[REDACTED]

Joan McMahon, Chair

[REDACTED]

Mary T. Withers

[REDACTED]

Steve Brannan

APPROVED:

[REDACTED]

Theodore G. Grove, Chair, Department of Speech
Communication

[REDACTED]

C. William Savery, Interim Vice Provost for Graduate
Studies and Research

DEDICATION

This thesis is dedicated to my grandmother, Violet Clark, who more than anyone else, encouraged me to finish it, and to my parents, Joan and Larry, who provided support along with the typewriter and correcting tape.

ACKNOWLEDGEMENTS

I would like to give my deepest thanks to Joan McMahon, my thesis and academic advisor, for all of her help during the past (dare I say it) six years. Her assistance and support, both during and outside of working hours, was invaluable to me and enabled me to complete this thesis at long last with minimal loss of sanity.

I would also like to thank Mary T. Withers who kindly joined by thesis committee when I needed another member.

Last, but not least, I want to thank Ted for not continually bugging me about how I should have used a computer for my thesis instead of typing it over and over again. It definitely was more work, but I guess I am just old fashioned. (How do you turn on a computer?)

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
CHAPTER	
I INTRODUCTION AND STATEMENT OF PURPOSE	1
Introduction	1
Statement of Purpose	2
II REVIEW OF THE LITERATURE	4
The Screening Process	4
Purpose and Description	
Characteristics of Screening	
Tests	
Screening Rationale	
Selection of Screening	
Candidates	
Screening Focus	
Language Screening Devices	9
Northwestern Syntax Screening	
Test (NSST)	
Bankson Language Screening Test	
(BLST)	
The Clinical Evaluation of Language	
Functions (CELF) Elementary	
Screening Test	12
Description	
Standardization Sample	
Score Interpretation	
Studies Using the CELF-S	
Rationale for Studying the CELF-S	17

III	METHODS AND PROCEDURES	18
	Subjects	18
	Instrumentation	18
	Procedures	19
	Hearing Screen CELF-S Administration	
	Scoring Procedures and Data Analysis .	20
	Scoring Procedures Data Analysis	
IV	RESULTS AND DISCUSSION	22
	Results	22
	Discussion	23
V	SUMMARY AND IMPLICATIONS	29
	Summary	29
	Implications	30
	Clinical Implications Research Implications	
	REFERENCES	33
	APPENDICES	
	A CONSENT FORM	37
	B CELF-S PROTOCOL	39
	C DIAGNOSTIC TESTS ADMINISTERED BY THE TEST SCHOOL'S SPEECH-LANGUAGE PATHOLOGIST .	44

LIST OF TABLES

TABLE		PAGE
I	Scores of Subjects Who Failed the Screening	25

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

INTRODUCTION

One of the responsibilities of the speech-language pathologist (SLP) is to identify children with language deficiencies (Mecham, 1979). Because assessing language abilities can be time-consuming, a quick and efficient procedure is needed for pinpointing children who need more in-depth evaluation. One of the more widely-used procedures to accomplish this task is screening. Because there are a considerable number of children in the schools, SLPs who work in these settings need to determine and use the most effective and efficient method for locating children with possible language problems. One language screening instrument that has been designed to accomplish this is the Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S) (Semel & Wiig, 1980).

The CELF-S was designed to assist professionals such as SLPs, psychologists, and educators in identifying elementary students with possible language disabilities (Semel & Wiig, 1980). It provides a measure for the

screening of language processing and production abilities of kindergarten through fifth graders. The purpose of the CELF-S is to identify children in need of more diagnostic evaluation, not to provide information about strengths, weaknesses, or degree of impairment. The scores obtained can also be used to identify the status of a child's performance when compared to his peers. The total testing time of the CELF-S is relatively short, 10 to 15 minutes, which makes it a time-efficient screening instrument.

STATEMENT OF PURPOSE

The purpose of this study was to determine the effectiveness of the CELF Elementary Screening Test in identifying those children, in a public school second grade setting, who were in need of more thorough diagnostic evaluation. More specifically, this study sought to determine whether the CELF-S identified the same children as those who were on the caseload of the school's SLP.

This investigation sought to answer the following questions:

1. What is the percentage of false negatives produced by the CELF-S (i.e., the number of children with speech and language disorders who were not detected by the screening test) as compared to those receiving services?

2. What is the percentage of false positives produced by the CELF-S (i.e., the number of children without language disorders who failed the screening test) as compared to those receiving services?

CHAPTER II

REVIEW OF THE LITERATURE

THE SCREENING PROCESS

Purpose and Description

The first level in the diagnosis of language disorders is the screening process (Maynard, 1973; Van Hattum, 1981). For the SLP in the schools, screening plays an important role in determining the adequacy of a child's language skills (Neidecker, 1980).

The primary purpose of screening is to determine whether or not a problem exists (Emerick & Hatten, 1974; Neidecker, 1980; Van Hattum, 1981). The screening process seeks to screen out children considered to have "normal" language while identifying those children who require more in-depth diagnostic evaluation (Somers & Hatton, 1985). Screening does not seek to determine what the problem is, only that a problem exists.

Because screening in the schools can involve a considerable number of students, the SLP must be able to use the most effective and efficient procedure for identifying those children with language deficiencies. It is important that the screening instrument allow the clinician to

rapidly pass over children with "normal" language while detecting those individuals with suspected problems (Carrow-Woolfolk & Lynch, 1982; Sommers & Hatton, 1985). According to Battle (1981), screening must be done with "...maximum expertise (but) with a minimum expenditure of time, energy, and money."

Characteristics of Screening Tests

There are several important characteristics that comprise a good screening test. First, the test should have proven validity (Launer & Lahey, 1981). In other words, the test should measure what it is intended to measure. Another characteristic of a good test is that it is reliable. This refers to the repeatability of the test, i.e., if given repeatedly, the test should give similar results each time. A screening instrument should also minimize the number of false negatives (those children who pass the screening, but are later found to be disordered) and false positives (those children who fail the screening, who later prove to have no problem) (Battle, 1981; Schwartz, 1983). Schwartz (1983) states, "Screening procedures should minimize the likelihood that any truly disordered child passes."

Screening Rationale

Several reasons for screening children have been cited. Tuomi (1978) suggests that screening, which identi-

fies children with language deficits, will ultimately prevent problems from occurring by eliminating conditions that may lead to these problems. Neidecker (1980) and Sommers & Hatton (1985) cite Public Law 94-142 (the Education for All Handicapped Children Act, 1975) as supplying another reason to screen. Public Law 94-142 guarantees handicapped children the right to a free, appropriate education in the most unrestrictive environment and requires that an identification process be implemented for those children who need assessment of their communication status. Neidecker (1980) states that one of the most widely-used procedures for this purpose is screening. A third reason for screening is offered by Wallach & Lee (1981). The reality of "too many children and not enough time" is well-known by many speech-language clinicians in the schools. If a clinician had to do detailed testing on a large number of children in order to identify those with language problems, it would take a great deal of time. Since some of the characteristics of screening are that it be quick and effective, it will allow for the examination of a large group over a shorter period of time (Schwartz, 1983), and this can help ease the problem of large numbers and limited time.

Selection of Screening Candidates

Sommers and Hatton (1985) note that there is lack of

agreement as to the grades and groups of children to be screened. Maynard (1973) states that screening can be used with those students who have been referred to the clinician by staff members or outside professionals and with those children who have received intervention in the past. Children who are new to the school are also candidates for screening. In addition to this latter group, Battle (1981) included the traditional method of screening full grades, especially kindergarten through third.

Battle (1981) wrote that there is support for the routine screening of all children in kindergarten and first grade due to the rapid development and change in their speech and language. According to Van Hattum (1981), learning a basic language is more easily accomplished when people are young. Thus, it is important that attention be paid to the communication abilities of young children in preschool and the early elementary years. According to Karmiloff-Smith (1977), "...researchers and clinicians alike have become increasingly aware of the role of language learning and the effect of early language disorders upon academic success." Dore (1979) recognizes the cognitive, linguistic, and communicative systems as intersecting types of knowledge that are yet distinct. Wallach and Lee (1981) noted that with the large amount of information available regarding the relationship between early language disorders and later academic success, early

screening programs play a significant role in the school setting. Screening early in a child's school career may not only help to identify those children who will need evaluation and possible early intervention, but may also supply information about children who may be at risk for difficulties in their school work. Schiefelbusch and Bricker (1981) and Miller (1983) have reported on successful early intervention programs which offer support for establishing identification (or screening) programs in the early grades.

Screening Focus

When assessing a child's language skills, the functions being examined are receptive (decoding) and expressive (encoding) verbal abilities (Maynard, 1973; Mecham, 1979; Neidecker, 1980; Stark, Tallal, & Mellits, 1982). Assessment of children with possible language delays or disorders should involve evaluation of the phonologic, semantic, morphological, and syntactic components of language (Neidecker, 1980; Stark, Tallal, & Mellits, 1982). A screening instrument provides a quick estimate of these areas, while a diagnostic instrument (used after screening) investigates each area in more depth and detail.

LANGUAGE SCREENING DEVICES

In the preliminary research for this study, this author found in 1988 that the Northwestern Syntax Screening Test (Lee, 1971), the Bankson Language Screening Test (Bankson, 1977), and the CELF-S were among the few screening tests on which research information could be found. There were, however, other screening instruments that were used at that time (as well as newer screening tests devised since that time) that are all worthy of use in future research.

Northwestern Syntax Screening Test (NSST)

One of the earliest-developed language screening instruments was the NSST (Lee, 1971). Like other screening tests, it does not intend to measure language skills, nor does it study syntax in-depth. Instead, it enables examiners to make a quick estimate of a child's syntactic development and to identify those children, between the ages of three and eight, who require more extensive evaluation of their syntactic abilities. It assesses receptive and expressive use of such syntactic forms as plurals, verb tenses, possessives, prepositions, and negatives, to name a few. The NSST is administered individually and takes approximately 20 minutes to complete. When interpreting the results, Lee (1971), advised that children be referred for further evaluation if their scores are more

than two standard deviations below the mean on the receptive or expressive sections or both.

Although the NSST has been widely used (Lee, 1977), it has been criticized as having norms that are geographically and socioeconomically narrow, which limits wide applicability.

Ratusnik, Klee, and Ratusnik (1980) developed a short form of the NSST that provides normative data on a much larger number of children, a wider geographical distribution, and a broader socioeconomic base. The authors state that the short form reduces the length of the original test by 45% (approximately 10 minutes) while retaining 95% of its discriminative capacity.

Bankson Language Screening Test (BLST)

The BLST (Bankson, 1977) was designed to assess a child's psycholinguistic and perceptual skills in a 25 minute time period. It consists of 17 subtests organized into five general categories: semantic knowledge, morphological rules, syntactic rules, visual perception, and auditory perception. Unlike the NSST (and other screening tests), the BLST is designed to evaluate expressive language specifically. Thus, scoring of the test focuses primarily on this aspect of language. According to Bankson (1977), expressive language was concentrated on for three reasons: 1) It is the mode of language through which most

people become aware of a language disorder; 2) Expressive language ultimately becomes the end goal of intervention; and 3) Tests of expressive language are not as available as receptive instruments. Due to the fact, however, that information about receptive abilities is also useful, provision is made in the test for supplemental receptive testing of certain items. They are not counted in the final score, however.

The BLST is administered individually to children ranging in age from 4.1 to 8.0 years and takes approximately 25 minutes to complete. It purports to discriminate between children with normal language abilities, those who are considered to be in need of further in-depth testing and children who will be most likely to receive language intervention in the future.

In a study by Blaxley, Clinker, and Warr-Leeper (1983), the BLST was found to be moderately accurate in identifying language impairments in kindergarten children who placed below the 10th percentile on the Developmental Sentence Scoring (Lee, 1974). The authors feel, however, that the length of the administration time (approximately 25 minutes) reduces the feasibility of using the BLST for large-scale screening.

THE CLINICAL EVALUATION OF LANGUAGE FUNCTIONS
(CELF) ELEMENTARY SCREENING TEST

Description

The CELF-S (Semel & Wiig, 1980) was designed to assist in identifying elementary students with potential language disabilities. Rather than pinpointing the specific strengths or weaknesses of language functions, the test's intent is to identify those children in need of more thorough evaluation. The authors' purpose for the CELF-S is to provide a measure for screening the language processing and production abilities of children in kindergarten through the fifth grade. (There is also an advanced form of the CELF-S for grades five through twelve.) The test is administered individually and takes approximately 10 to 15 minutes to complete.

The screening test consists of two sections. The first contains items that test a child's ability to process language. These 31 items were selected in order to probe aspects of the following: 1) Accuracy in phoneme discrimination; 2) Sentence formation rules (morphology and syntax); 3) Interpretation of words and logical relationships among sentence components and linguistic concepts; and 4) Retention and recall of word and action sequences. This first section follows a "Simon-Says"-type format. To administer this section of the test, the examiner should have the child stand directly in front of

him/her, at least three feet away. The child's responses are scored as either correct or incorrect, and a score of 1 or 0 is given respectively.

The second section of this screening test deals with language production. The 17 items included in this section are designed to probe features of the following processes: 1) Agility and accuracy in phoneme production; 2) Ability to recall, identify, and retrieve words and concepts; 3) Accuracy in serial recall; and 4) Immediate recall of model sentences. Here, spoken stimuli are presented which require a spoken response. In this section, the child's responses should be recorded on the score form verbatim. As in the processing section, the responses are scored as correct or incorrect with a 1 or 0 given accordingly.

Standardization Sample

The CELF-S was standardized on 1,346 normal children in grades kindergarten through fifth. While it was intended that the standardization sample include an equal number of males and females at each grade level, the percentage within specific grades varied between 44% and 56%. The subjects were taken from four major geographic regions with limited information regarding socioeconomic status (SES). Semel and Wiig (1980) state that because the most important factor is to select students representative of

their school, the SES status of the sample does not need to be questioned. The racio-ethnic background of the subjects was close to a 1980 Census report for the population of children between the ages of 5 and 14. Grade level, rather than age-level, was used in the normative sample. The authors felt that because language growth and school learning appear to be related, using grade levels would be more in line with the purpose of a screening test, i.e., to identify those children with language abilities below expectations for their grade level.

Score Interpretation

The CELF-S provides a raw score and percentile rank, by grade level, for an overall total score as well as for the processing and production sections individually. The authors suggest that, until local norms can be established, those children with a total percentile rank below 15, or a percentile rank lower than 10 on either the processing or production sections, be referred for more in-depth evaluation.

Studies Using the CELF-S

There have been few research studies that have used the CELF Elementary Screening Test. Ribner, Becker, Marks, Kahn, and Wolfson (1983) conducted a study using the CELF-S in order to determine its appropriateness and validity for a school population in New York City. Speech

and language personnel in the New York school system filled out questionnaires on all special education children who had previously been given a language evaluation. On the questionnaire, each child's deficit was rated on a 5-point scale, ranging from 1, no language deficit to 5, a severe language deficit requiring a total communication program. Each child was then administered the CELF-S. Findings from the CELF-S were compared with the ratings made by the speech and language specialists. Results of this study indicated that the classification of a child's performance on the CELF-S did not show much agreement with the classifications based on the language specialists' recommendations even when the most optimal grouping of subjects based on the total CELF-S test was used. The percentage of this agreement was 63.1. The authors concluded that, based on this inability to agree, the effectiveness of the CELF-S as a screening instrument is questionable. The authors, however, do state that one could argue that a lack of substantial agreement between the CELF-S and the language specialists' ratings may not be a result of a deficiency in the CELF-S but rather due to the categories derived from the ratings of the specialists. The authors responded to this, however, by stating that since there was considerable variability between the scores obtained in this study and those listed in the CELF-S' standardization sample, the discriminatory powers of the CELF-S are

lessened.

Illerbrun, Haines, and Greenough (1985) used the CELF Elementary Screening Test in a study that attempted to obtain an estimate of predictive validity for five language screening tests in classifying children relative to their performance on three "combined" diagnostic criterion measures. The five screening tests included the Kindergarten Language Screening Test (KLST) (Gauthier & Madison, 1978), the Bankson Language Screening Test (BLST) (Bankson, 1977), Fluharty Preschool Speech and Language Screening Test (Fluharty, 1978), CELF Elementary Screening Test (Semel & Wiig, 1980), and the Language Identification Screening Test for Kindergarten (LIST-K) (Illerbrun, McLeod, Greenough, & Haines, 1984). The diagnostic measures consisted of the Test of Language Development (Newcomer & Hammill, 1977), the Test of Auditory Comprehension of Language (Carrow, 1973), and the Carrow Elicited Language Inventory (Carrow, 1974). The five screening tests were administered to each child during the first two and three months of their kindergarten year. Three months later, the children were administered each of the three diagnostic tests. Results of the study found that all but the KLST were highly valid and reliable, with the LIST-K being the most efficient as a mass kindergarten language screening test.

RATIONALE FOR STUDYING THE CELF-S

One of the reasons the CELF-S was selected for this study was because of the minimal amount of research done on this instrument. This test appeared to have all of the characteristics of a good screening test, but this could not be supported due to the paucity of research.

Since this study was begun in 1988, a revised version of the CELF-S has been developed. This author, however, was unable to obtain a copy of this revised test, and as a result, was unable to determine the differences between the old and new CELF-S.

CHAPTER III

METHODS AND PROCEDURES

SUBJECTS

The subjects for this study were 50 second grade children attending a public elementary school in Hillsboro, Oregon, who were enrolled in a regular education classroom. Each subject passed a puretone audiometric screening in order to be included in the study. The subjects' sex, socioeconomic status, and intelligence level were not controlled because this study (and the CELF-S) sought to identify all children with language problems regardless of these factors. More information about students and scores could have been gained if these had been considered. (See Appendix A.)

INSTRUMENTATION

A portable Beltone audiometer ANSI 1969 was used to conduct the audiometric screening of the subjects.

The Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S) (Semel & Wiig, 1980) is a screening instrument designed to measure the language processing and production abilities of children in kindergar-

ten through fifth grade. It consists of 31 oral commands and 17 expressive items, which do not follow a strict order of progression. (See Appendix B.) The examiner administered the entire test to each of the subjects following the instructions given in the test manual. The testing required approximately 10 to 15 minutes per subject.

PROCEDURES

Hearing Screen

The examiner brought each child individually into a quiet and well-lit testing room in the school. Rapport was established through conversation on the way to the testing room. Upon entering the room, the children were seated in a chair, with their backs to the examiner, and administered the audiometric puretone screening. The clinician presented puretones at 500, 1000, 2000, and 4000 Hz through earphones at 25 dB in the right ear (in the left ear if the right failed). The subjects had to pass the screening in one ear, and only those who passed the screening were included in this research study.

CELF-S Administration

Upon passing the hearing screening, the children were directed to stand three to four feet in front of the examiner. The processing section of the CELF-S was then administered following the directions printed in the test

manual. No materials other than the CELF-S protocol were used. The protocol was attached to a clipboard and held on the examiner's lap, away from the children's line of vision.

After completing the first section of the CELF-S, the children were seated at a table in a chair placed to the left of the examiner. The production portion was then administered. Throughout the testing procedure, the examiner reinforced the general behaviors of the children, such as good listening and paying attention. Upon completion of testing, the children were reinforced for their cooperation with a sticker and verbal praise. The total screening time, including the audiometric screening, took approximately 15 to 20 minutes per child.

SCORING PROCEDURES AND DATA ANALYSIS

Scoring Procedures

The examiner recorded each child's responses on the test protocol, following the instructions for scoring in the test manual. Correct responses were scored as a 1, while incorrect responses received a score of 0. Once the screening was completed, the points were totalled and a raw score obtained for the overall test, as well as for each subtest. Each child's percentile rank was then determined and recorded on the score form.

Data Analysis

The data obtained from the screening was analyzed descriptively. Below the 15th percentile was considered failing for the total test score while below the 10th percentile was considered failing for the processing and production subtests. Those students whose scores fell below the 15th percentile for a total test score or below the 10th percentile on one or both of the subtests were identified as "screened in" (positives). Those children whose scores placed at the cutoffs or above were identified as "screened out" (negatives). The results of the screening were then compared with the names of children on the SLP's caseload, and the percentage of false positives and false negatives were determined.

CHAPTER IV

RESULTS AND DISCUSSION

RESULTS

The purpose of this study was to determine the effectiveness of the CELF-S in identifying those children, in a public school second grade, who were in need of more thorough diagnostic evaluation as a result of failing this language screening test. The names of the children found to have failed the CELF-S in this study were compared to those students on the school SLP's caseload in order to determine whether the CELF-S identified all of the second graders currently being seen for language intervention. Of the 50 students screened in this study, four children scored below the CELF-S' cutoff percentiles, thus failing the screening. The scores of these four students provide the data base for answering the following questions.

The first question posed by this study was: What is the percentage of false negatives produced by the CELF-S as compared to those previously identified by the school's SLP? In this study, of the 50 children screened, the three children already on the SLP's caseload were also identified by the CELF-S. The percentage of false nega-

tives produced by the CELF-S, therefore, was 0%.

The second question this investigation sought to answer was: What is the percentage of false positives produced by the CELF-S? Results of this study revealed that of the 50 children screened, only one false positive was produced by the screening test, meaning that this student had not been previously evaluated and identified as language delayed but had failed the screening. The percentage of false positives produced by the CELF-S, therefore, was 2%.

DISCUSSION

This study sought to determine whether the CELF-S was effective in identifying those children in a public school second grade who needed more in-depth language evaluations. The results obtained from this investigation appear to indicate that the CELF-S was effective in this identification.

Of the 50 children screened, only one false positive was produced by the CELF-S. Further in-depth evaluation to determine whether a language delay was present in this subject could not be undertaken, however, because of time limitations. A possible mitigating factor that could have affected this particular student's performance was that the child was Hispanic and used English as a second language. The student did not qualify for the district's

Language Transition Room (LTR), however, in which children of migrant families and those for whom English is not the primary language are given instruction. The eligibility criteria for this room is the following: A student must not be able to speak any English or a student must obtain a score of 3 or below on the Language Assessment Scales (an instrument designed to assess the language competency of Spanish speakers in both Spanish and English).

Because the CELF-S seeks to identify all children in a regular second grade classroom with possible language problems, no effort was made to omit those children who were not completely competent with the English language and who did not qualify for the LTR. There appears to be a valid concern about the appropriateness of using this screening instrument with this student since the CELF-S was standardized on more than 76% white children at the elementary level and less than 6% of children of Spanish origin and also because there was no way of knowing the student's English competency for this study. The authors of the CELF-S reported an attempt to omit those test items that would be biased against a child of a different race. They stated, however, that because the sample of non-white children was relatively small, a statistical analysis might not detect a significance. Semel and Wiig (1980) did caution, therefore, about interpreting scores of culturally and ethnically different children on the basis

of their published norms for this test because the examiner might not know how the differences in other cultures may contribute to scores obtained on the screening test.

In this study, the percentile rankings of the subjects' scores were studied because those subjects with total percentile rankings below 15 or a processing or production percentile rank of less than 10 were recommended as possibly needing further in-depth evaluation. No other scoring criteria were used. Table I shows the scores for the four children who failed the screening.

TABLE I
SCORES OF SUBJECTS WHO FAILED THE SCREENING

Subject	Sex	Total Test		Processing		Production	
		Raw	%ile	Raw	%ile	Raw	%ile
5	F	24	15	20	28	4	5
27	M	24	15	15	9	9	23
37	F	25	17	15	9	10	32
41	F	20	8	17	15	3	4

As shown in Table I, two of the four children failed the CELF-S because they fell below the cutoff criteria for the production subtest. One of these children (#5) showed a 23 percentile point difference between the processing and production scores which could indicate a problem in

expressive language skills. The other student who failed the production subtest (#41) also fell below the total test percentile ranking of less than 15. This resulted because the child had also received a lower score on the processing subtest, although the processing score was not low enough to be failing. This could indicate below average receptive skills and a delay in expressive language skills.

A third student (#37) failed the CELF-S because a failing score was obtained on the processing subtest. The child's score of 9 on this subtest was 23 percentile points below the score on the production subtest, which is somewhat unusual as comprehension is widely viewed as preceding production in development (Bloom & Lahey, 1978). Dale (1976) provides a possible explanation for this when he states that there appear to be instances in which certain parts of language are produced appropriately yet comprehension is not yet mastered.

The fourth subject (#27) was the Hispanic student. He failed the CELF-S because his processing percentile fell below the cutoff. His total test percentile was only one point above the cutoff, indicating the production percentile was also low, yet not below the cutoff score. It is difficult, if not impossible, however, to determine if these scores truly indicate a need for further in-depth language evaluation. As previously mentioned, the authors

of the CELF-S caution users of the test in interpreting scores of culturally and ethnically different children on the basis of their published norms for this test. One reason for this, according to Semel and Wiig (1980), is because the percentages of these students in the norming sample did not always equal the proportion of such children in the total population. Also, scores of these children are questionable when there are differences in a child's cultural or ethnic background that may contribute to the scores in ways that the examiner or the test itself cannot precisely measure. As a result, it is difficult to determine if any deficits really exist.

Although not a part of this study, it is interesting to note the relationship between the effectiveness of teacher referrals and the accuracy of the CELF-S in identifying those children needing further testing. As no second grade referrals were made by classroom teachers during the school year this study took place, teacher referral effectiveness was inferred from past referrals.

During the school year that this study was conducted, the SLP had received no second grade referrals from the classroom teachers. Of the three second graders who were already identified as language disordered at the time of this study, only one was known to have been referred by the classroom teacher in first grade. The other two language delayed children had moved into the district already

identified as having language problems. According to the SLP's past records, six children (who were second graders at the time of this study) had been referred for speech-language evaluations when they were in first grade. Of these six children who had been referred by their first grade teachers and had been given in-depth evaluations, only two had qualified for speech-language services. Of these two children, one had moved out of the district the previous year, and the other was one of the second graders on the SLP's caseload who had failed the CELF-S during this study.

When looking at the four students who failed the CELF-S, only one (just mentioned) was known to have been referred for language testing (in first grade). Two of the four had moved into the district already identified as language delayed while the fourth had never been referred or identified (the Hispanic student).

Based on this limited information, it does not seem that teacher referral was a very accurate method of identifying those children who needed to be screened; however, since there was a small subject sample and data was not very complete, this is a very cautious statement.

CHAPTER V

SUMMARY AND IMPLICATIONS

SUMMARY

One of the more widely-used methods for pinpointing children in need of more in-depth language evaluation is screening. One language screening instrument designed to accomplish this in an effective and efficient way was the Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S) (Semel & Wiig, 1980).

The purpose of this study was to determine the effectiveness of the CELF-S in identifying those children in a second grade setting, who were in need of more thorough evaluation. This study sought to answer the following questions: 1) What is the percentage of false negatives produced by the CELF-S?, and 2) What is the percentage of false positives produced by the CELF-S?

Fifty second graders from a public elementary school were selected as subjects. Each subject passed a puretone audiometric screening and had received parental permission to be in the study. The subjects' sex, socioeconomic status, and intelligence were not controlled since this investigation sought to identify all children with

possible language problems, regardless of these other factors.

Of the 50 children screened, the three students who had previously been identified as language impaired by the school's speech-language pathologist (SLP) were also identified by the CELF-S. The percentage of false negatives, therefore, was 0%. Only one student out of the 50 screened (who had not been previously tested or identified as language impaired) failed the screening. The percentage of false positives produced by the CELF-S, therefore, was 2%.

IMPLICATIONS

Clinical Implications

The CELF-S has been a screening tool used by SLPs. The results of this study support the use of this instrument for second grade students who speak Standard English as their primary language. It is the opinion of this researcher that anyone who uses the CELF-S should be cautious when using it to screen students for whom English is not the primary language. Although a foreign student might not qualify for a special class such as a language transition room or English as a Second Language (ESL) classroom, it is this investigator's opinion that one cannot eliminate the possibility that a language barrier problem could be the cause of failing test scores rather than

a language delay.

Speech-language pathologists who may want to use this screening instrument might consider collecting their own local norms since there may be school districts that do not match up well with the norming sample used by the CELF-S. Semel and Wiig (1980) recommend establishing local norms as they also realize that there may be school populations that do not compare as readily with their norming sample.

Research Implications

If a similar study is pursued, it is recommended that further in-depth evaluation be completed on students who fail the CELF-S and who are not already identified as language delayed in order to further determine the screening test's effectiveness.

This study appeared to produce valid results for the second grade. It is recommended that similar research be conducted on other grade levels to determine if the CELF-S results are as valid as they were for the second graders tested.

One limitation of this study was the small number of students already identified as language delayed who were used to determine if the CELF-S identified the appropriate children. As a result, it is difficult to extrapolate the results of this study to a larger group. If a similar

study is conducted in the future, it is recommended that a much larger data base be used.

Further examination of the effectiveness of teacher referral and the CELF-S needs to be conducted, as this was not a focus of this study.

REFERENCES

- Bankson, N.W. (1977). Bankson Language Screening Test. Baltimore: University Park Press.
- Battle, D.E. (1981). Speech screening. Seminars in Speech, Language, and Hearing, 2, 37-51.
- Berry-Luterman, L., & Bar, A. (1971). The diagnostic significance of sentence repetition for language-impaired children. Journal of Speech and Hearing Disorders, 36, 29-39.
- Blaxley, L., Clinker, M., & Warr-Leeper, G. (1983). Performance on two screening tests compared with performance on DSS. Language, Speech, and Hearing Services in Schools, 14, 38-46.
- Bloom, L., & Lahey, M. (1978). Language development and language disorders. New York: John Wiley & Sons.
- Boehm, A. (1967). Boehm Test of Basic Concepts Manual. New York: The Psychological Corporation.
- Boehm, A. (1971). Boehm Test of Basic Concepts. New York: The Psychological Corporation.
- Boehm, A. (1986). Boehm Test of Basic Concepts-Revised. San Antonio, TX: The Psychological Corporation.
- Bright, H. (1973). BTBC-A new tool for assessing language concepts. Language, Speech, and Hearing Services in Schools, 4, 140-142.
- Carrow, E. (1973). Test for Auditory Comprehension of Language. Austin, TX: Learning Concepts.
- Carrow, E. (1974). Carrow Elicited Language Inventory. Austin, TX: Learning Concepts.
- Carrow-Woolfolk, E., & Lynch, J. (1982). An integrative approach to language disorders in children. New York: Grune & Stratton.

- Clark, E. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T.E. Moore (Ed.), Cognitive development and the development of language. New York: Academic Press.
- Dale, P.S. (1976). Language development: Structure and function. New York: Holt, Rinehart, & Winston.
- Dore, J. (1979). What's so conceptual about the acquisition of linguistic structures. Child, 6, 129-137.
- Dunn, L.M., & Dunn, L.M. (1981). Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service.
- Emerick, L.L., & Hatten, J.T. (1974). Diagnosis and evaluation in speech pathology. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
- England, G.F., Gray, B.B., & Ryan, B.P. (1971). The Programmed Conditioned Language Test. Monterey, CA: Accelerated Achievements Associates, Inc.
- Fluharty, N.B. (1978). Fluharty Preschool Speech and Language Screening Test. New York: Teaching Resources.
- Gauthier, S.V., & Madison, C.L. (1978). Kindergarten Language Screening Test. Tigard, OR: C.C. Publications.
- Illerbrun, D., McLeod, J., Greenough, P., & Haines, L. (1984). Language Identification Screening Test for Kindergarten. Saskatoon, Saskatchewan: EDEXC Occasional Publications, University of Saskatchewan.
- Illerbrun, D., Haines, L., & Greenough, P. (1985). Language identification screening test for kindergarten: A comparison with four screening and three diagnostic language tests. Language, Speech, and Hearing Services in Schools, 16, 280-292.
- Karmiloff-Smith, A. (1977). More about the same: Children's understanding of post articles. Child Language, 4, 377-394.
- Launer, P.B., & Lahey, M. (1981). Passages: From the fifties to the eighties in language assessment. Topics in Language Disorders, 1, 11-29.

- Lee, L. (1971). Northwestern Syntax Screening Test. Evanston, IL: Northwestern University Press
- Lee, L. (1974). Developmental Sentence Analysis. Evanston, IL: Northwestern University Press.
- Lee, L. (1977). Reply to Arndt and Byrne. Journal of Speech and Hearing Disorders, 42, 323-327.
- Lucas, E.V. (1980). Semantic and pragmatic language disorders: Assessment and remediation. Rockville, MD: Aspen Systems.
- Maynard, J. (1973). Task force report on school speech, hearing, and language screening procedures. Language, Speech, and Hearing Services in Schools, 4, 109-119.
- Mecham, M.J. (1979). Testing procedures appropriate for speech pathology. In B.B. Hutchinson, M.L. Hanson, and M.J. Mecham (Eds.), Diagnostic handbook of speech pathology. Baltimore: The Williams & Wilkins Company.
- Miller, J.F. (1983). Identifying children with language disorders and describing their language performance. In J.F. Miller, D.E. Yoder, and R. Schiefelbusch (Eds.), Contemporary issues in language intervention. Rockville, MD: ASHA (The American Speech-Language-Hearing Association).
- Neidecker, E.A. (1980). School programs in speech-language: Organization and management. Englewood, N.J.: Prentice-Hall, Inc.
- Newcomer, P.L., & Hammill, D.D. (1977). The Test of Language Development: Test Manual. Austin, TX: Empiric Press.
- Newcomer, P.L., & Hammill, D.D. (1982). The Test of Language Development-Primary. Austin, TX: PRO-ED.
- Ratusnik, D.C., Klee, T.M., & Ratusnik, C.M. (1980). Northwestern Syntax Screening Test: A short form. Journal of Speech and Hearing Disorders, 45, 200-207.

- Ribner, S., Becker, L., Marks, S., Kahn, P., & Wolfson, F. (1983). A validation study of the elementary and advanced screening tests of the Clinical Evaluation of Language Functions (CELF). Language, Speech, and Hearing Services in Schools, 14, 215-222.
- Schiefelbusch, R., & Bricker, D. (1981). Early language intervention. Baltimore: University Press.
- Schwartz, R.G. (1983). Diagnosis of speech sound disorders in children. In I.J. Meitus and B. Weinberg (Eds.), Diagnosis in speech-language pathology. Baltimore: University Park Press.
- Semel, E.M., & Wiig, E.H. (1980). Clinical Evaluation of Language Functions-Elementary Screening Test. San Antonio, TX: The Psychological Corporation.
- Sommers, R.K., & Hatton, M.E. (1985). Establishing the therapy program: Case finding, case selection, and caseload. In R.J. Van Hattum (Ed.), Organization of speech-language services in schools. San Diego: College-Hill Press.
- Stark, R.E., Tallal, P., & Mellits, E.D. (1982). Quantification of language abilities in children. In N.J. Lass (Ed.), Speech and language: Advances in basic research and practice. New York: Academic Press.
- Tuomi, S. (1978). Screening of speech and language disorders: International perspective. Paper presented to world congress on the future special education. First, Stirling, Scotland.
- Van Hattum, R.J. (1981). Diagnosis of communication disorders. Seminars in Speech, Language, and Hearing, 2, 1-9.
- Wallach, G.P., & Lee, A.D. (1981). Language screening in the schools. Language, Speech, and Hearing Services in Schools, 2, 53-75.

APPENDIX A

CONSENT FORM

CONSENT FORM

I agree to let my child, _____, participate as a subject in the study entitled "A Retrospective Study of the Clinical Evaluation of Language Functions Elementary Screening Test (CELF-S)." This study will be conducted by Tami Caldwell under the supervision of Joan McMahon, thesis director, Speech and Hearing Sciences Program, Portland State University.

In this study, my child will be given a brief hearing screening and the CELF-S which involves following instructions given in a "Simon Says" format and also involves completing sentences, answering questions, and repeating different words and sentences. There are no risks involved in this study. My child is free to refuse to participate or to withdraw his/her prior approval without prejudice. I, too, am free to refuse to let my child participate or to withdraw him/her from the study without prejudice. In order to insure my child's anonymity, no names will be used when results are tabulated and presented. Instead, he or she will be assigned a number, which will be used for identification purposes. The test will require approximately fifteen to twenty minutes to complete.

In the event that my child is identified by the CELF-S as needing further language testing, and my child has not been tested by the school's speech-language pathologist, a further language test may be administered in order to check the accuracy of the CELF-S. This language test would be the Test of Language Development-Primary (TOLD-P) which involves pointing to pictures, defining words, imitating sentences, and completing sentences. If results of this test indicate a language problem, the examiner (Tami Caldwell) will notify me by telephone. This test would take approximately 45 minutes to complete.

The purpose of this study is to determine the effectiveness of the CELF-S in identifying those children in a second grade setting, who are in need of more thorough in-depth language evaluation.

Signature of Parent

Date

Signature of Child

(A child over 7 years of age must be in the study and must sign his name)

Please complete and have your child return this form to the classroom teacher tomorrow or as soon as possible. (I have enclosed 2 copies of this consent for one for you to keep for your records, and one for you to sign and return.) If you have any questions, please feel free to contact me at 544-5775. If your child experiences any problems that are the result of participation in this study, please contact the secretary of the Human Subjects Research and Review Committee, Office of Grants and Contracts, 303 Cramer Hall, Portland State University, 464-3417. Thank you for your cooperation.

Tami Caldwell

APPENDIX B

CELF-S PROTOCOL



Clinical Evaluation of Language Functions

Honor Alesing-Semel
Elizabeth H. Wiig

**ELEMENTARY LEVEL
SCREENING TEST**

Name _____ Date _____

Birthdate _____ Grade _____ Sex _____

School _____ Teacher _____

Home Address _____

Other Relevant Background _____

Summary and Comments _____

Examiner _____

	Total	Processing	Production
Raw Score:			
Percentile Rank:			



THE PSYCHOLOGICAL CORPORATION
HARCOURT BRACE JOVANOVICH, INC.

Language Processing Screening Items

DEMONSTRATION ITEMS.

1. Simon says: Touch your ear.
2. Simon says: Touch your toes.
3. Point to your nose.
4. Simon says: Clap your hands, wave good-bye, salute the flag.

TRIAL ITEMS.

1. Simon says: Hold up your pinkie.
2. Touch your knees.
3. Simon says: Touch your mouth, point to your shoe.

TEST ITEMS. (Each command may be read only once.)

SCORING: Record the child's responses by marking through the appropriate score (1 or 0).
Correct responses after the first reading of an item score one point. Errors score zero points.

STIMULUS	RESPONSE	CORRECT	INCORRECT
1. Simon says: Touch your hand, touch your head.		1	0
2. Simon says: Point to your wrist.		1	0
3. Simon says: Point to your toes.		1	0
4. Simon says: Point to the lowest part of your face.		1	0
5. Simon says: Clap your hands slowly.		1	0
6. Raise your hands above your head quickly.		1	0
7. Simon says: Point to all of your fingers except your thumbs.		1	0
8. Point to one of your feet.		1	0
9. Simon says: Point to your eyes.		1	0
10. Simon says: Touch your knees, touch your toes, touch your nose.		1	0
11. Simon says: After I say the word "clap," you clap your hands. (PAUSE) Tap, snap, clap, slap.		1	0
12. Put your hands in front of your face.		1	0
13. Simon says: Put your hands up, put your hands down.		1	0
14. Simon says: Touch your head above your ears.		1	0
15. Simon says: Raise your left knee, touch your nose.		1	0
16. Simon says: Touch your hip.		1	0
17. Touch your ear, touch your thumb.		1	0
18. Simon says: Point to your longest finger.		1	0
19. Simon says: When I say the word "nose," touch your nose. (PAUSE) Toes, knees, eyes, nose, hands.		1	0
20. Simon says: Put your hands between your knees.		1	0
21. Simon says: Touch your leg below the knee.		1	0
22. Simon says: Point to all of these: hands, hips, head.		1	0
23. Touch your elbow.		1	0
24. Simon says: Clap your hands, tap your forehead, snap your fingers.		1	0
25. Simon says: Turn right, then face me.		1	0
26. Simon says: Point to your cheek, chin, chest.		1	0
27. Turn around.		1	0
28. Simon says: If I say the words "Raise your hand," then do it. Listen carefully. Raise your foot. Raise your knee. Raise your hand. Raise your elbow.		1	0
29. Simon says: Point to some of these: knees, nose, ear, eye, toes.		1	0

STIMULUS	RESPONSE	CONNECT	INCORRECT
30. Simon says: Put your right hand on your right hip, your left hand on your left shoulder, your right hand on your left hip.		1	0
31. Simon says: Turn to the left, then face me.		1	0

Language Production Screening Items

TRIAL ITEMS.

- Count to five.
- Repeat this word after me: HIPPOPOTAMUS.
- Complete this phrase: "On my feet I wear socks and _____."

TEST ITEMS. (Each command may be read only once.)

SCORING. Record the child's responses verbatim and score each by marking through the appropriate score (1 or 0).

STIMULUS	RESPONSE	CONNECT	INCORRECT
1. Complete this phrase: "Red, white, and _____."		1	0
2. Complete this phrase: "Knife, fork, and _____."		1	0
3. Tell me the names of the days of the week. (Sun. Mon. Tue. Wed. Thurs. Fri. Sat.)		1	0
4. Tell me everything you can about orange juice. (Allow 60 second period for responding.)	NO. OF DISCRETE FEATURES NAMED = _____ (Score 1 if 3 or more; score 0 if fewer than 3.)	1	0
5. Tell me which month comes after March.		1	0
6. Tell me the letters of the alphabet. (a b c d e f g h i j k l m n o p q r s t u v w x y z)		1	0
7. Repeat this sentence after me: "Jack likes hamburgers with ketchup."		1	0
8. Repeat this word after me: "com <u>pl</u> yishment." kŭm plŭ' ish munt		1	0
9. Repeat this sentence after me: "Jack likes hamburgers with ketchup and mustard."		1	0
10. Repeat this word after me: "ac <u>n</u> apheminopia." tāk' ə fē mīn ɔ̄ pē ə		1	0
11. Count to twenty by twos. (2, 4, 6, 8, 10, 12, 14, 16, 18, 20)		1	0
12. Tell me the three letters that come after "X."		1	0
13. What is the opposite of "full"?		1	0

STIMULUS	RESPONSE	CORRECT	INCORRECT
14. What is the opposite of "add"?		1	0
15. Repeat this sentence after me: "Jack likes hamburgers with relish, mustard, and ketchup."		1	0
16. Repeat this sentence after me: "Pale luminous feelings blithely painted the ocean."		1	0
17. Repeat this sentence after me: "Jack likes french fries and hamburgers with ketchup, onions, mustard, and relish."		1	0

APPENDIX C

DIAGNOSTIC TESTS ADMINISTERED BY THE TEST SCHOOL'S
SPEECH-LANGUAGE PATHOLOGIST

Test of Language Development-Primary (TOLD-P)

The TOLD-P (Newcomer & Hammill, 1982) is a revised edition of an earlier test by Newcomer and Hammill (1977), the Test of Language Development. The TOLD-P is designed to offer both a method of comparing a child's spoken language skills with those of the child's peers as well as a method of comparing a child's own abilities across different areas. The TOLD-P is a receptive/expressive measure that assesses different features (syntax, semantics, and phonology) and systems (listening and speaking). The TOLD-P is individually administered to children ages 4.0-8.11 years and takes from 30 minutes to one hour to complete.

Programmed Conditioned Language Test (PCLT)

The PCLT (England, Gray, & Ryan, 1979) is an instrument designed to measure the number of expressive language errors made by a child when imitating sentences. Research has found that sentence imitation is a valuable tool when assessing a child's linguistic development (Berry-Luterman & Bar, 1971; Dale, 1976). The PCLT contains 55 stimulus items (assessing both question forms and regular sentence structures) which are then imitated by the child being tested. This test is administered individually and takes approximately 20 to 40 minutes to complete.

Boehm Test of Basic Concepts-Revised (BTBC-R)

The BTBC-R (Boehm, 1986) is a revised edition of the original Boehm Test of Basic Concepts (Boehm, 1971). It is designed to assess the mastery of 50 concepts that are considered essential to understanding and communication and which are related to achievement in the first few years of school (Boehm, 1967). The purpose of the BTBC-R is both diagnostic and remedial (Bright, 1973). It attempts to identify individual children whose overall level of concept mastery is low as well as pinpointing specific concepts not known by a child (Boehm, 1971). The BTBC-R is designed for use with grades kindergarten through third and can be administered individually or to an entire classroom of children. It takes approximately 15 to 20 minutes to administer.

Peabody Picture Vocabulary Test-Revised (PPVT-R)

The assessment of vocabulary development is widely viewed as an important part of an overall language assessment battery (Bloom & Lahey, 1978; Clark, 1973; Dale, 1976; Lucas, 1980). The PPVT-R (Dunn & Dunn, 1981) was one instrument designed for this purpose. It is a non-verbal test that measures a person's receptive vocabulary for Standard American English. The PPVT-R consists of two comparable forms, L and M, each containing 175 pictorial items arranged in order of increasing difficulty. Each

item contains four illustrations from which the subject is to select the picture best illustrating the meaning of the stimulus word. The PPVT-R is individually administered and takes approximately 10 to 20 minutes to complete. It was designed for use with people ages 2½ to 40 years.