EXAMINER'S MANUAL
SCREENER VERSION

DRAFT

Mabel L. Rice, Ph.D.
Kenneth Wexler, Ph.D.

Printed in the United States of America

0158896629

1 2 3 4 5 6 7 8 9 10 11 12 A B C D E

Acknowledgments

The Rice/Wexler Test of Early Grammatical Impairment grew out of a program of research directed by the authors and funded by awards from the National Institute of Deafness and Other Communication Disorders (RO1 DC01803), and research directed by Rice and also funded by NIDCD (P50 DC02746 with Bruce Tomblin as the principal investigator).

The real beginning was the youngsters with language impairments I encountered during my early years as a speech-language pathologist, especially those for whom there was no apparent cause for their impairment. My frustration with the limitations I encountered with the means of identification available for affected children led me to a course of inquiry that has played out over the following decades. I spent the fall of 1990 in Ken Wexler's lab at MIT, where one could learn about the emerging models of morphosyntax and engage in ongoing tutorials with Ken and a marvelous group of doctoral students, including: David Poeppel, William Snyder, Sergey Avrutin, and Colin Phillips; and later expanded to Carson Schutze and Jenny Ganger, along with faculty members Alec Marantz and David Pesetsky. I express my deep appreciation to Ken, for his interest in my questions about children with unexplained difficulties in grammatical acquisition, and for the cohesion and rigor of the theory of Optional Infinitives in children. This theory allowed for the formulation of precise predictions to be evaluated in children with Specific Language Impairment (SLI), and, perhaps, a better understanding of how some of the pieces of the puzzle are configured and how they interface with other pieces. The model ultimately led to an instrument that may help us to better identify young children with limited language competency.

Investigation of children with SLI, and the necessary comparison groups of unaffected children, is an extensive empirical endeavor that involves the contributions of many people. I wish to thank Karla Barnhill for her steadfast and highly valuable contributions as a coordinator of data collection, and for her firsthand, astute observations of many of the children who participated in the experimental protocols. I am grateful to Janna Oetting, Pat Cleave, and Sean Redmond for their suggestions and advice; for their participation in the development and evaluation of new elicitation tasks during their participation in the lab; for their contributions to data collection; and for their creative extensions of the lab work in their dissertations and subsequent programs of independent research. Mary Howe and Hiromi Morikawa provided data summaries and documentation of an ongoing series of outcomes. Janet Marquis has been there throughout for queries about quantitative analyses and general guidance on matters statistical; she brought her expertise to the formal analyses of the experimental data. Scott Hershberger carried part of the quantitative load for the first reports of growth curve data analyses. Pam Hadley provided valuable procedural suggestions early on, along with insightful interpretive extensions in her dissertation study and her subsequent investigations of young children. My deep appreciation goes to Patsy Woods for her daily contributions as Administrative Assistant; she functions as a real Earth Mother to the entire team. In addition to these people, there are many other students and staff who have participated and who have my appreciation for their contributions to the lab.

Beyond the Kansas site, I wish to express my appreciation to Bruce Tomblin for his invitation to join the collaboration on Specific Language Impairment for studies on the large samples of children recruited and studied at the University of Iowa. This invitation allowed for the collection of data on some of the experimental probes with a wider group of language impaired children, and valuable collaborative interactions

with my scientific colleagues Jeff Murray, Lawrence Leonard, Susan Ellis Weismer, Carol Miller, Robert Kail, Marc Fey, and Lawrence Shriberg. A special thanks goes to my colleague down the hall, Hugh Catts.

I thank Carolyn Mervis for her invitation to collaborate with her on studies of children with Williams Syndrome; Martha Crago, Fred Genesee, and Joanne Paradis for their invitation to collaborate with them on studies of French-speaking children with SLI; and Helen Tager-Flusberg and Jeanne Roberts for their invitation to collaborate with them on studies of children with autism. I thank Kate Taylor and Steve Zubrick for their invitation to spend time developing new lines of inquiry while in residence in Perth, Australia.

A number of people within the community of scholars have helped shape the direction of my thinking about grammatical morphology and language impairments, but of course they are not responsible for my continued limitations. They include Jill and Peter deVillers, Judith Johnston, Harold Clausen, Dorothy Bishop, Celia Jakubowicz, Steven Crain, Jeannette Schaeffer, Yonata Levy, Gina Conti-Ramsden, Cornelia Hamann, Roz Thornton, and Andrew Radford. I wish also to thank Judith Cooper for her wise consultations on grant matters.

The making of this test was a 3-year endeavor. Thanks to the many people at The Psychological Corporation for their efforts in the development of the Rice/Wexler Test of Early Grammatical Impairment: Dr. Charles "Chuck" Wilkins and Dr. Agnes Stephenson provided valuable quantitative assistance; the text benefited from the capable editing of Dawn Dunleavy, Pam Parmer, and Michelle Girard. Dr. Carol Waryas was supportive from the very beginning. Jan Laurent, Project Manager, coordinated and oversaw the entire project, and did so with impressive competency and grace. I also thank the individuals who participated as bias reviewers and the many professionals who participated as field examiners.

To my friends Kathleen, Kim, Susan C., Susan K., Marilyn, Toni, Scott, Rhonda, and others—thanks for your support. For my daughter, Melinda, and for Toshi, with love. For the children with language impairments and those who wish to teach them.

Mabel Rice
July 2001

This test is based on long years of research that Mabel Rice and I conducted on Specific Language Impairment (SLI); I thank Mabel for introducing me to the fascinating, difficult, and important problem of language impairment and for all the patience and courage that she showed in allowing ideas and results that I believed in, but that were so different from received views on impairment, to enter into this problem. The NIH supported this SLI research; I hope that they will feel that their faith was justified.

The knowledge that we now have of SLI could not have begun to develop if there hadn't been long years of difficult, if exciting, research on esoteric scientific problems involving the Optional Infinitive stage in many languages. I will never forget what I have elsewhere called the "great community of OI researchers"—the large group of people around the world who jumped into the problem, collecting and analyzing data, pointing out counterexamples in language X, confirming data in language Y, and running different analyses; finding loopholes, inconsistencies, and even, occasionally, praise.

This field has grown so quickly that even I, an optimist, am stunned. If this test benefits children, which is my dearest hope, there are many people to be thanked—so many people involved from around the world that I dare not try to list them all. Let me just mention the following:

Among students at MIT who contributed so much to move me in the right direction on early development; Amy Pierce Brand, David Poeppel, Carson Schutze, Colin Phillips, Sergey Avrutin, Jenny Ganger, Masha Babyonyshev, William Snyder, Tania Ionin; postdoctoral fellows and lab visitors Teresa Guasti, Roz Thornton, Frank Wijnen, Astrid Ferdinand, Vicenc Torrens, Manuela Schoenenberger, Thomas Lee, Jeannete Schaeffer, Judy Baek, and Hyeon-jin Lee; Juergen Weissenborn, a lab visitor who heard my lectures, listened carefully, and found evidence in his own French and German data; Bonnie Schwartz, a lab visitor who let me understand how second language acquisition differs; my knowledgeable, argumentative, and interested faculty colleagues, David Pesetsky and Alec Marantz; Nina Hyams, the person who kept generative approaches to early inflectional development alive for so many years, was always available to give me the benefit of her expertise; Rita Manzini dropped in occasionally with the trenchant remark that proved so useful; Corneilia Hamann always had new challenges to offer; Bernhard Rohrbacher pushed forward in analyzing data from a very young child; Celia Jakubowicz, Lea Nash, and Harald Clahsen prodded me with different ideas; Steven Crain, while not working on this particular issue, was always a comrade in figuring out how to discover what kids were capable of in language; Luigi Rizzi who had the scientific insight and foresight to get interested in the problem early and invited me to teach in Geneva, thereby opening up not only his own intriguing work on the problem, but a whole community's.

This is just the beginning; I am leaving out too many people and not even attempting to list all the researchers on SLI who eventually reacted and contributed in various ways. The point is that, while the scientific research that laid the foundation for the Rice/Wexler Test of Early Grammatical Impairment may be sometimes situated in papers with several names on it, in essence it is the product of a great scientific community.

It is also the product of the support that I continue to receive from loved ones. I could not have managed to maintain such effort without the love and support of my wife, Sherry Wexler (who was deeply involved in her own doctoral work during this period), Paul, and Stephanie and Chris. I thank them from the bottom of my heart.

Ken Wexler
July 2001

Table of Contents

Chapter 1 Overview 1

Purpose 1

Theoretical Background:
A Morphosyntactic View of Children's Grammar 1

A Clinical Marker Approach to Language Assessment 4

Advantages of Using Rice/Wexler Compared to
Other Language Measures 6

Chapter 2 Administration and Scoring 7

Probes 7

Test Components 7

Examiner's Manual 7

Stimulus Manual 7

Record Form 7

Training Videotape 7

Manipulatives 8

User Qualifications 8

Administration Time 8

Using the Screening Test 8

Scores Reported 8

Testing Considerations and Procedures 9

Test Environment 9

Establishing a Rapport 9

Parents/Caregivers 10

The Testing Session 10

Completing the Record Form 10

Computing Chronological Age 10

Administration Procedures 11

General Directions 11

Practice and Trial Items 12

Using Prompts 12

Phonological Probe 12

Directions 13

Scoring Responses and Marking the Record Form 13

Completing the Phonological Probe Summary 14

Third Person Singular Probe 14

Directions 14

Recording Responses 15

Scoring Responses and Marking the Record Form 15

Completing the Third Person Singular Probe Summary 16

Past Tense Probe 16
Directions 16
Recording Responses 18
Scoring Responses and Marking the Record Form 18
Completing the Past Tense Probe Summary 19
Supplemental Scoring for the Past Tense Probe 20
*Be/Do Probe . . . . .
Directions . . . . .
Recording Responses . . . . .
Scoring Responses and Marking the Record Form . . . . .
Completing the Be/Do Probe Score Summary . . . . .
*Grammaticality Judgment Probe . . . . .
Directions . . . . .
Scoring Responses and Marking the Record Form . . . . .
Completing the Grammaticality Judgment Probe Summary . . . . .
Completing the Summary Scores 20
*Calculating Elicited Grammar Composite 20
Calculating the Screening Score 20
Identifying Criterion Scores 22
Marking the Growth Curves 22

Chapter 3 Interpretation 25
Scores Calculated 25
About the Scores 25
Development of Criterion Scores 26
Using the Rice/Wexler in Clinical Practice 27
Establishing Eligibility for Services 28
Using Rice/Wexler as a Screening Tool 28
Interpreting the Rice/Wexler Scores 29
Elicited Grammar Composite 29
Interpreting the Phonological Probe Results 29
Interpreting the Third Person Singular Probe Results 30
Interpreting the Past Tense Probe Results 31
*Interpreting the Be/Do Probe Results . . . . .
*Interpreting Grammaticality Judgment Probe Results . . . . .
Production Probes and A' Values: Implications for General Intervention Strategies 34
Comparisons With Related Test Data 35
Interpreting Children's Performance Relative to Nonverbal Intelligence or Parent Education Level 36

Chapter 4 Research and Development 37
History of the Instrument 37
Probe Development and Previous Testing 37
Tryout Research 40
Bias Review 41
Examiner Input and Feedback 44
Analyses 44
Standardization Research 44
Selection and Qualification of Examiners 45
Description of the Standardization Sample 45
Scores 50
Development of Criterion Scores 53
Growth Curves 53
Previous Studies and Comparisons to Rice/Wexler Outcomes 55

Chapter 5 Technical Characteristics 69
Reliability 69
Test-Retest Reliability 69
Validity 71
Evidence Based on Test Content 71
Evidence Based on Internal Structure 71
Evidence Based on Relations to Other Variables 74
Convergent Evidence 75
Discriminant Evidence 77
Relationship Between Rice/Wexler and Parent Education Levels 81

Appendix A Criterion Scores 83

Appendix B Sensitivity and Specificity 85

Appendix C Probe Score Look-up Tables 101

***Appendix D A' Look-up Tables** . . . . .

Appendix E Scoring Examples 103

***Appendix F Rice/Wexler Test of Early Grammatical Impairment Examiners** . . . . .

References 107

***This material is listed for reference only, but can be found in the full version of this manual at the following link: http://www2.ku.edu/~cldp/MabelRice/pubslist.shtml**

List of Figures

Figure 1.1 The Bell-Shaped Distribution 5
Figure 1.2 Distribution of Children's Performance in the Area of Grammar 6
Figure 2.1 Calculating Chronological Age 11
Figure 2.2 Recording and Scoring Responses for the Third Person Singular Probe 17
Figure 2.3 Recording and Scoring Responses for the Past Tense Probe 21
*Figure 2.4 Examples of Editing Be/Do Responses . . . . .
*Figure 2.5 Recording and Scoring Responses for the Be/Do Probe . . . . .
*Figure 2.6 Recording and Scoring Responses for the Grammaticality Judgment Probe . . . . .
Figure 2.7 Completing the Summary Scores Sections 23
Figure 4.1 Box and Whiskers Plot for Third Person Singular 52
Figure 4.2 Box and Whiskers Plot for Past Tense 52
*Figure 4.3 Box and Whiskers Plot for Be/Do Score (Be) . . . . .
*Figure 4.4 Box and Whiskers Plot for Be/Do Score (Do) . . . . .
*Figure 4.5 Box and Whiskers Plot for Elicited Grammar Composite . . . . .
*Figure 4.6 Box and Whiskers Plot for Grammaticality Judgment Dropped Marker . . . . .
*Figure 4.7 Box and Whiskers Plot for Grammaticality Judgment Agreement . . . . .
*Figure 4.8 Box and Whiskers Plot for Grammaticality Judgment Dropped -ing . . . . .
*Figure 4.9 A' Calculations for Grammaticality Judgment Probe—Normal Language Group . . . . .
*Figure 4.10.A' Calculations for Grammaticality Judgment Probe—Language Disorder Group . . . . .
Figure 4.11 Past Tense Calculations—Normal Language Group 66
Figure 4.12 Past Tense Calculations—Language Disorder Group 66

List of Tables

Table 2.1 Examples of Responses to Reprompt and to Not Reprompt 12
Table 4.1 Tryout Research Sample by Age and Language Status 41
Table 4.2 Rice/Wexler Bias Panel Members 43
Table 4.3 Rice/Wexler Standardization Sample by Age—Children in the Normal Language Group 46
Table 4.4 Rice/Wexler Standardization Sample by Age—Children in the Language Disorder Group 47
Table 4.5 Rice/Wexler Standardization Sample by Gender—Children in the Normal Language Group 47
Table 4.6 Rice/Wexler Standardization Sample by Gender—Children in the Language Disorder Group 47
Table 4.7 Rice/Wexler Standardization Sample by Race/Ethnicity—Children in the Normal Language Group 48

Table 4.8 Rice/Wexler Standardization Sample by Race/Ethnicity—Children in the Language Disorder Group 48
Table 4.9 Rice/Wexler Standardization Sample by Region—Children in the Normal Language Group 48
Table 4.10 Rice/Wexler Standardization Sample by Region—Children in the Language Disorder Group 48
Table 4.11 Rice/Wexler Standardization Sample by Parent Education Level—Children in the Normal Language Group 49
Table 4.12 Rice/Wexler Standardization Sample by Parent Education Level—Children in the Language Disorder Group 49
Table 4.13 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Normal Language Group 51
Table 4.14 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Language Disorder Group. 51
Table 4.15 25th Percentile, Mean Score, and 75th Percentile for each Rice/Wexler Probe and for the Elicited Grammar Composite—Children in the Normal Language Group 54
Table 4.16 Analysis of Variance for Rice/Wexler Probes and for the Elicited Grammar Composite 57
Table 4.17 t test for the Rice/Wexler Probe Scores and for the Elicited Grammar Composite. 58
Table 4.18 Repeated Measures ANOVA—Group by Age and A' Type. 61
Table 4.19 Repeated Measures ANOVA Table—Normal Language Group and Language Disorder Group 62
Table 4.20 t test between Rice/Wexler Grammaticality Judgment Probe Scores 63
Table 4.21 t test between Rice/Wexler Grammaticality Judgment Probe Scores—Children in the Language Disorder Group 64
Table 5.1 Test-Retest Means, Standard Deviations, Mean Absolute Score Differences, and Stability Coefficients for Rice/Wexler Probes and Elicited Grammar Composite (n = 106). 70
Table 5.2 Correlations of the Rice/Wexler probes and Elicited Grammar Composite—Children in the Normal Language Group 72
Table 5.3 Correlations of the Rice/Wexler Probes and Elicited Grammar Composite—Children with Language Impairments. 73
Table 5.4 Means, Standard Deviations, and Correlations Between CELF®–Preschool Word Structure and Rice/Wexler Probes and Elicited Grammar Composite 76
Table 5.5 Means, Standard Deviations, and Correlations Between CELF®–3 Word Structure and Rice/Wexler Probes and Elicited Grammar Composite 77
Table 5.6 Means, Standard Deviations, and Correlations Between CELF®–Preschool Basic Concepts and Rice/Wexler Probes and Elicited Grammar Composite 78

Table 5.7 Means, Standard Deviations, and Correlations Between CELF®–3 Word Classes and Rice/Wexler probes and Elicited Grammar Composite for Children with Language Disorders. 79

Table 5.8 Means, Standard Deviations, and Correlations for the Elicited Grammar Composite and Non-Verbal IQ Scores for Children with Language Disorders 80

Table 5.9 Means, Standard Deviations, and Correlations for the Elicited Grammar Composite and Grammaticality Judgments and Parent Education Levels 81

*This material is listed for reference only, but can be found in the full version of this manual at the following link: http://www2.ku.edu/~cldp/MabelRice/pubslist.shtml

1 Overview

Purpose

The *Rice/Wexler Test of Early Grammatical Impairment* (Rice/Wexler) is an individually administered clinical tool that can be used for the identification, diagnosis, screening, and follow-up evaluation of grammatical deficits in young children, ages 3 through 8 years, who speak Standard American English (SAE). The test focuses on a particular area of grammar that is known to be difficult for young children with language impairment at the age that this grammatical competence becomes well established in children with normal language acquisition. Poor performance on finiteness marking as measured in the Rice/Wexler can serve as a clinical marker that identifies "affected" children, or children who have impairments in this area of grammar. Early identification of language impairments in young children, especially children without other disabilities (sometimes known as Specific Language Impairment), is critical to ensuring the early intervention required to prepare children for the language demands of the early school years. In combination with assessments of vocabulary and speech development, the Rice/Wexler can provide documentation of fundamental elements of language for children during the preschool years and the school entry and early elementary school years.

Theoretical Background: A Morphosyntactic View of Children's Grammar

The *Rice/Wexler Test of Early Grammatical Impairment* is well motivated theoretically and empirically. The test focuses on the area of grammar in the domain of *finite verb morphology*. In English, morphemes that mark finiteness include third person singular *-s*, regular and irregular past tense, copular and auxiliary forms of *Be,* and auxiliary forms of *Do*. Finite verb morphology is recognized as a particularly promising area for measures that are sensitive to factors that indicate diagnosis of Specific Language Impairment (SLI) (Tager-Flusberg & Cooper, 1999) in young children.

Clinical language assessment instruments have traditionally followed a general framework that includes assessing a wide range of language competencies in the areas of semantics (vocabulary), syntax (word order), morphology (e.g., noun or verb affixes), comparative adjectives (e.g., "good," "better," "best"), or pragmatics (e.g., narratives or story comprehension). In contrast, the Rice/Wexler focuses on grammatical morphology. Although multiple areas of language may be affected in children who have language impairments, grammatical

morphology has obligatory properties that enhance the clinician's ability to identify affected children.

Children's acquisition of grammatical morphemes has been an object of scientific interest for several decades. Since the 1950s, investigators have observed that very young children tend to drop some of the obligatory grammatical morphemes of English; for example, children may say, "*dog running," when what they seem to mean to say is, "the dog is running," or they say, "*dog walk home," when what they seem to mean to say is, "the dog walks home." Following new developments in theoretical linguistics, studies began to appear in the late 1980s that treated young children's morphology as part of their syntactic system. The term "morpho-syntax" was increasingly used to recognize this relationship and to highlight that the underlying linguistic knowledge was not limited to that of a lexical stem-plus-affix (push + -ed). The morphosyntactic perspective presented a way for investigators to see that although young English-speaking children sometimes, though not always, drop morphemes, these children also show that they understand the relationship that exists between morpheme use and syntax. Children do not, for example, place inflected verbs in places in the sentence where inflected verbs cannot appear. In the examples that follow (and throughout the manual), the asterisk (*) indicates an ungrammatical sentence: "*Runs the dog home" or "*the dog walks not." These sorts of errors are rarely, if ever, seen in children's utterances.

Kenneth Wexler (1992, 1994, 1996) named the early stage of children's morphosyntax an "Optional Infinitive" stage. The "infinitive" part of the label comes from Wexler's studies of languages, such as French or German, where infinitival forms of the verb appear in the surface phonology as affixes. Wexler's studies (which support other studies on many languages) reported that children sometimes use infinitival forms of verbs in places in their sentences where conjugated, or finite, forms of verbs are expected. This suggests that the children regarded the infinitival form as an optional form of a verb, hence the term "Optional Infinitive." This stage also can be thought of as "Optional Finiteness."

The morphosyntactic perspective, with its emphasis on finiteness and rules for word order, has brought an important new approach to English morphology, with two new insights put forth by Wexler (1992, 1994) that potentially may have significant relevance in identifying children with language impairments. The first insight is that a small set of morphemes share the same underlying grammatical property of finiteness, even though their surface properties are different. The morphemes included in the following set share this underlying linguistic property of finiteness marking. They are illustrated here in simple clauses that require the finiteness marker (shown in italics) for the clause to be grammatically well-formed.

1. Patsy walks home. (third person singular subject, present tense)
2. Patsy walked home yesterday. (no subject agreement, regular past tense)
3. Patsy ran home yesterday. (no subject agreement, irregular past tense)
4. Patsy is walking. (third person singular subject, auxiliary present tense)
5. Patsy is happy. (third person singular subject, copular present tense)
6. Does Patsy walk home? (third person singular, auxiliary present tense)

Wexler's second insight is that in those English contexts where children use uninflected forms of English verbs, such as "*Patsy run home" instead of "Patsy ran home," the word run may function as a nonfinite form, as though to the child the nonfinite and finite forms were optional variations in that sentence context. In sentences such

as 1–6, then, the italicized finiteness markers may be dropped, as illustrated in 7–12, where the omitted forms appear in parentheses.

7. *Patsy walk(*s*) home.
8. *Patsy walk(*ed*) home yesterday.
9. *Patsy run(*ran*) home yesterday.
10. *Patsy (*is*) walking.
11. *Patsy (*is*) happy.
12. *(*Does*) Patsy walk home?

This may lead to dropped affixes on verbs, and to dropped forms of auxiliary and copula Be or auxiliary Do, because those grammatical forms exist to meet the requirement that every main clause in English have a finiteness marker (although in some contexts this marker is silent, such as first person present tense, e.g., "I walk home").

To summarize, the Optional Infinitive account of the young child's grammar generates three very powerful predictions:

- A small set of finiteness markers should be interrelated in a child's acquisition of grammar.
- An English-speaking child sometimes may drop these markers in sentences.
- Morphemes that do not share the finiteness-marking function do not follow the same path of change over time as morphemes with this linguistic property.

Note that within the morphosyntactic perspective it is important that a complete sentence structure is present to establish the obligatory context for the morpheme use. This means, for example, that in English clauses a subject must be present to establish the need for third person singular present tense. In elliptical contexts an infinitival form is allowed; for example, if someone asks "What does Patsy like to do?" a speaker can reply "walk home" and the finite marker -s is not required.

Within this perspective, the morphophonological properties of finiteness markers are viewed as different elements in their ease of acquisition; for example, in English, -s appears on nouns to mark regular plurals (e.g., cats), and on verbs to mark third person singular present tense (e.g., walks). Even though the morphemes share similar phonological rules, their underlying linguistic functions are different. So young children may be very accurate in their use of regular plurals, while at the same time inconsistent in their use of third person singular present tense -s in obligatory contexts.

The Optional Infinitive/morphosyntactic model provides a linguistically enriched view of Brown's (1973) 14 morphemes. Brown selected this group of morphemes primarily because it is possible to identify contexts in which the use of a given form is obligatory. It is possible to calculate the percentage correct in obligatory contexts and to use that percentage as an index of change as a child's grammar moves toward the adult form. Brown selected progressive -ing (as in "Patsy is running"), the prepositions in and on, plural -s, regular and irregular past tense, third person present singular -s, the articles a and the, and copular and auxiliary Be. This set of morphemes has since been included in many omnibus language assessment instruments to evaluate children's grammatical development. Note that, in English, the progressive -ing does not mark finiteness, but indicates the aspect distinction that an action is ongoing. In the example clause, "Patsy is running," the auxiliary Be, which can be present or past tense, carries the grammatical function of finiteness marking and must appear for

the clause to be grammatically correct. Thus, finiteness is not always marked by affixes to the verb stem and not all verbal affixes mark finiteness.

The morphosyntactic perspective emphasizes a morpheme set that marks finiteness, and the obligatory nature of finiteness for a grammatically correct clause (as constrained by rules for word order), rather than the rules involved in adding an affix or an internal vowel shift to a verb stem. At the level of assessing a child's grammar, this perspective focuses attention on those morphemes that appear in obligatory contexts and that serve the same underlying linguistic function. It enables us to see, for example, that the auxiliary Do form, although not included in Brown's original 1973 study, should be examined in addition to the Be forms.

Although a number of studies examining acquisition of morphology by young children were completed by the early 1990s, there was at that time no careful, empirical evidence that documented the longitudinal course of the acquisition of finiteness markers by English-speaking children. Nor had when children shift from optional use of these markers to the obligatory use of the adult grammar been examined. In short, there was no clear evidence regarding how long the period of "child grammar" persists in children without language impairments, although there were strong indications that children start with an inconsistent use of the finiteness markers.

A Clinical Marker Approach to Language Assessment

The notion of a clinical marker comes from the medical literature and relates to the observation that a particular symptom of affectedness can be especially accurate in the identification of individuals who are affected with a particular condition or impairment. A good marker is defined as one that not only is likely to identify affected individuals, but also to identify unaffected individuals.

Mabel Rice and Kenneth Wexler have collaborated to explore the area of finiteness marking as a clinical marker of language impairments in children. In theory, if accurate use of finiteness marking is regarded as optional for normally developing children for a portion of their early development, it may be that it is more difficult for children with language impairments. If this is the case, it follows that affected children (children with language disorders) could be detected by evaluating their performance on particular grammatical structures. A test that assesses finiteness marking, then, would be a valuable tool both in clinical use and in the investigation of factors that contribute to language impairment. Research into the genetic basis of language impairments, for example, requires precise methods of identifying affected individuals (see Rice, 2000, for further discussion of these issues).

Conventional language assessments assume that the language aptitude of a given sample of children (e.g., all 5-year-olds) is distributed in a pattern similar to a bell-shaped curve (see Figure 1.1). Individuals are thought to be distributed along a range of performance levels, so that a few people display very high values (along the right-hand side of the curve), a few people display very low values (along the left-hand side of the curve), and most people (about 68%) score in the middle. Individuals with language impairments are those who fall at the bottom end of the normal distribution of language competence.

This approach has led to many important research findings and clinical applications; at the same time, however, it has some significant limitations. The first limitation is that there is no intrinsic criterion for where to draw the line between "normal" and "disordered" performance. This leads to a certain inevitable arbitrariness and an ongoing debate

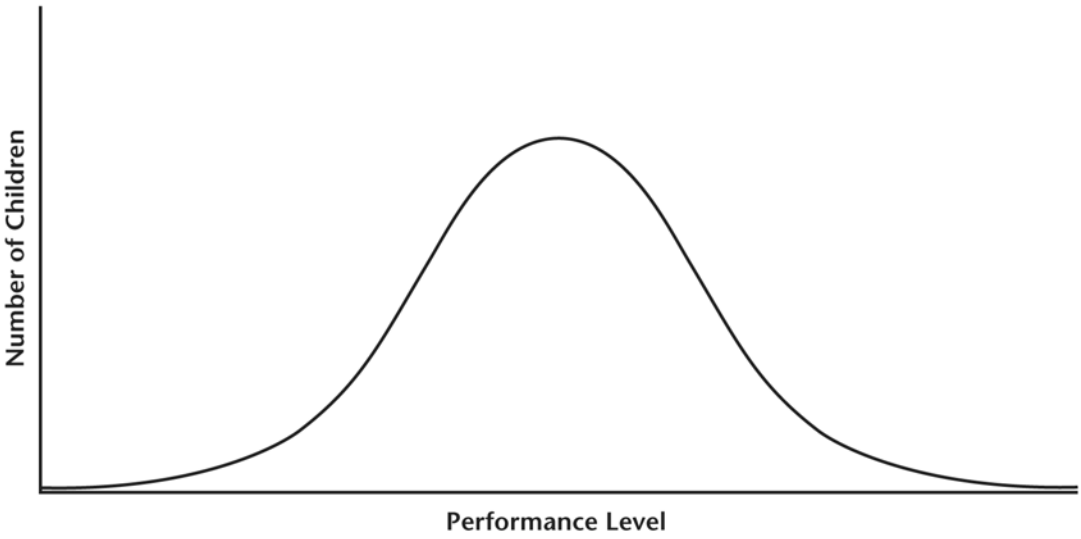


Figure 1.1 The Bell-Shaped Distribution

about the "best" clinical cut-off level. A second limitation is that there is no obvious way to interpret a test score in terms of particular linguistic content. Because the tests are typically constructed across multiple dimensions of language and grammatical functions, they provide no provisions regarding how a child's performance on particular linguistic competencies can be understood in terms of which areas are and which areas are not affected. A third limitation is that it is not possible to interpret a child's performance relative to the expected adult model of language, or a child's level of progress toward the adult level. In other words, the child's position on the normal curve does not tell us exactly what the child knows about grammar.

The fundamental rules of grammar, such as those that govern finiteness-marking in English, do not allow a bell-shaped distribution across individuals. Consider the examples in items 1–6 (page 2) as compared to items 7–12 (page 3). Speakers of Standard American English know that the clauses in items 1–6 are well-formed (correct or grammatical) and the clauses in items 7–12 are not (incorrect or ungrammatical). Once normally developing children pass through the optional phase, they have generally mastered each of these aspects of grammar with consistency that approximates that of adult users. Grammar users are not distributed as in a bell curve; instead they bunch up at the top end of the distribution because they know all or most of these grammatical principles.

We know that by kindergarten age most typically developing children know the basic properties of grammar. We know also that children do not show these properties in their first simple sentences. Instead, for a period of time children differ from adults in that they tend to generate sentences such as those in items 7–12. Over time, however, they begin to generate items such as 1–6 and are unlikely to continue to generate items such as 7–12. To identify children with language impairments we must find those children who continue to generate sentences such as 7–12 at a frequency greater than expected for a particular age level.

Consider 5-year-old children: it is expected that most of them will perform like adults in this part of grammar. Figure 1.2 shows an hypothesized distribution of children, with respect to grammar. The distribution bunches at the right hand side of high performance level, at or near the adult level of grammatical accuracy. When we look at the properties of language children know in the adult form, we restrict the variation across unaffected (normally developing) children; instead of about 68% of them performing in the middle range, most of them will be performing at the high end

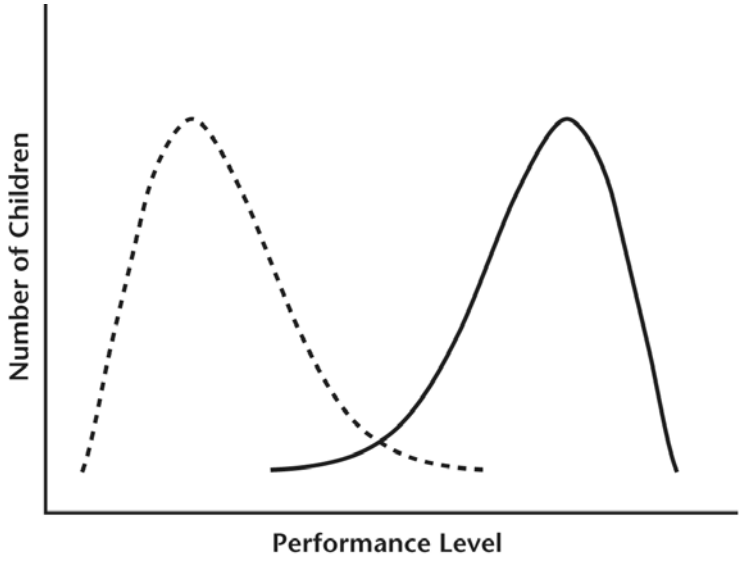


Figure 1.2 Distribution of Children's Performance in the Area of Grammar

of accuracy. Children with language impairments who do not know this part of the grammar by 5 years of age would cluster at the low end of the performance range. With this type of distinction it is easier to identify children who do not know the properties of grammar (affected or disordered children) because they cluster below the unaffected children.

Advantages of Using Rice/Wexler Compared to Other Language Measures

In addition to the robust psychometric properties of the test that are described, there are multiple advantages to using the Rice/Wexler for the assessment of children in the 3 through 8 years age range.

1. Its focus on finiteness is conceptually sound in terms of the linguistic properties of adult grammar.
2. Performance on the test can be directly interpreted as describing fundamental properties of what a child knows about grammar.
3. Performance on the test can be interpreted in terms of a child's progress toward the adult grammar.
4. The test focuses on a property of English grammar that is known to be well mastered by children before they enter school.
5. It focuses on a property of grammar that is known to be difficult for children with language impairments.
6. It can identify affected children whose sole developmental deficit is language impairment, such as children with Specific Language Impairment (SLI).
7. It is well suited to identify children of school-entry age who need early intervention.
8. It has high levels of sensitivity and specificity, leading to accurate identification of affected children, without a high rate of false identification of unaffected children.

Administration and Scoring

2

Probes

There are five sections, or Probes, in the Rice/Wexler Test of Early Grammatical Impairment:

- Phonological Probe
- Third Person Singular Probe
- Past Tense Probe
- Be/Do Probe
- Grammaticality Judgment Probe

Test Components

The *Rice/Wexler Test of Early Grammatical Impairment* is composed of an Examiner's Manual, a Stimulus Manual, a Record Form, a Training Videotape, and several manipulatives.

Examiner's Manual

The Examiner's Manual contains detailed information about administration, scoring, and interpretation procedures. It also contains information about research, test development, and technical qualities.

Stimulus Manual

The Stimulus Manual contains the color picture stimuli necessary to administer test items.

Record Form

The Record Form contains abbreviated directions for administering, recording, and scoring the test items; the scripts for the Be/Do Probe and Grammaticality Judgment Probe; and all trial and test items. The first page of the Record Form includes a place to record demographic information and a Summary Scores section.

Training Videotape

The Training Videotape is designed to assist you in learning to administer the test. The videotaped demonstration supplements the information presented in the Examiner's Manual and on the Record Form.

Manipulatives

The child's interaction and participation in the test is facilitated by using items that are included in the test kit. These items are: three stuffed teddy bears (one of each: white, tan, brown); one stuffed cat; one stuffed bug; one rabbit hand puppet; one plastic googly-eyed finger puppet; one plastic drinking glass; three smile-face plastic creatures (three different colors), referred to in the test as "moon guys" or "robots"; one yellow cloth; two plastic spoons; three plastic forks; one milk carton; one juice carton; one plastic apple; and one plastic hamburger. In addition to what is included in your test kit, you will need to provide one shoe-sized box and one box of tissues.

User Qualifications

The *Rice/Wexler Test of Early Grammatical Impairment* was designed to be administered by professionals knowledgeable in conducting and interpreting language assessments and evaluations. Individuals such as speech-language pathologists, educational diagnosticians, psychologists, and early childhood educators may be appropriate to administer this instrument if they have the necessary training and experience in diagnosing language disorders in young children. Individuals using this instrument should also have experience in administering individual assessments and in interpreting the results.

Administration Time

Administration of the *Rice/Wexler Test of Early Grammatical Impairment* takes approximately 45 minutes, depending on the child's age.

Using the Screening Test

An abbreviated form of the Rice/Wexler can be used as a screener to quickly determine whether or not a child needs additional services. To conduct a screening using the instrument, you will first administer the Phonological Probe to determine if the child is able to produce the phonemes being tested. Next, you will administer and score the Third Person Singular Probe followed by the Past Tense Probe, and then compare the scores obtained from these probes with the Screening Test criterion score for the child's age. Criterion scores for the screening test can be found in Appendix A**.** It will take about 10 minutes to complete the screening test.

Using the Screening Test function can be very useful when you have to make a determination as to whether or not a child needs further evaluation, or when you have to determine a child's developmental status or school readiness. Because the administration is quick and easy, using the screening test is a valuable tool to use for large scale screening endeavors.

Scores Reported

Probe scores are calculated for each probe and for the Elicited Grammar Composite. Criterion scores are provided for each half-year age level between the ages of 3 years 0 months and 6 years 11 months and for each year level for children 7 and 8 (see Appendix A). In addition, growth curves are provided that show the performance

of children in the normal language group relative to their acquisition of the adult grammar. The growth curves for each probe appear on the *Record Form* and a discussion regarding how to interpret these data is included in Chapter 3.

Testing Considerations and Procedures

It is extremely important that you read and understand the instructions for administering and scoring the *Rice/Wexler Test of Early Grammatical Impairment* before you administer it to a child. Follow all instructions precisely to maintain test reliability and to make appropriate interpretations based on test results. In addition to following standard testing procedures, keep in mind other variables that may influence the child's performance, such as the testing environment, the rapport you establish with the child, the reinforcement and encouragement you offer, and the appropriateness and timing of breaks during the testing session.

The Rice/Wexler was standardized on, and criterion scores were developed based on, a sample of children who were all administered the test in the same manner. It is important that you administer the instrument in this same fashion, otherwise the established criterion scores will have little applicability for the child you are testing.

Test Environment

Conduct testing in a quiet, well-lit and well-ventilated room, removed from distractions and disruptions. Have the child sit at a table with you, positioned so the child is comfortable and can easily see you and interact with you. Have the manipulatives near by so you can access them easily, but placed so they will not distract the child. Have the Record Form on the table but out of the child's direct view, so you can mark his or her responses easily. For younger children, you should be flexible with the seating arrangement (e.g., some children may work better on the floor, or seated in the lap of a parent/caregiver, than seated at a table). Use your clinical judgment to determine the most appropriate arrangement to ensure the most effective testing results.

Establishing a Rapport

Before you begin testing, spend some time with the child to get acquainted and to establish a rapport. It is important that the child be comfortable in the setting before testing begins. During test administration, you may provide general comments or reinforcing statements such as, "I like the way you are working" or "We're almost done." If the child appears reluctant to respond, offer encouragement by saying, "Give it a try" or "It's okay if you are not sure." Such comments may be necessary to maintain the child's attention and motivation, especially for younger children. Do not tell the child if his or her responses to test items are correct or how many items he or she answered correctly. (You may provide additional prompts and/or correct responses for the practice or trial items only.)

Note: Be cautious in using the manipulatives to establish a rapport with the child. Because the manipulatives are not used during the Rice/Wexler test administration until the last two probes, early introduction of the toys may cause some children to become distracted, making it difficult for you to gain and maintain their attention.

Parents/Caregivers

Parents or caregivers may be present during testing, if necessary, to encourage the child's participation (this may be especially helpful for younger children). If parents or caregivers are present during testing, instruct them to not prompt or comment on the child's performance at any time during testing.

The Testing Session

Administer the entire test in one session, if possible. However, you may take additional sessions if necessary. If you need additional sessions to complete test administration, you should complete the entire administration within seven days of the first session.

During test administration, monitor the child's performance and behavior to ensure that he or she is attending well enough to complete testing, or to identify if he or she is becoming fatigued. If the child tires during testing, or becomes inattentive or non-compliant, discontinue testing. Use your clinical judgment to determine whether or not to continue the testing session or if it is more appropriate to discontinue testing and schedule an additional session.

If a short break is required, either for you or the child (e.g., getting a drink of water, a restroom break), or a rest period during testing seems necessary, schedule a break so that it does not interrupt administration of a probe. Do not interrupt testing during the middle of a probe unless absolutely necessary. If an emergency requires you to interrupt testing during a probe, use your clinical judgment to determine if you should continue testing where you left off, or if re-administering the entire probe is more appropriate. Consider both how many items had been administered before the break and how long the break was when making your decision about where to resume testing. If you have only one or two items left in the probe and the break is very brief, it makes sense to resume testing where you left off. However, if you have administered half the items and then you have to break for an hour, it would be best to start over with the practice items and instructions.

Completing the *Record Form*

Before you begin testing, complete the demographic information requested on page 1 of the *Record Form*. This section includes space for the child's name, gender, school, grade (if applicable), classroom teacher's name, and the examiner's name. Also enter the assessment date (first session), and the child's birth date. Compute the child's chronological age (see Figure 2.1) and write it in the box labeled Chronological Age.

Computing Chronological Age

Calculate the child's chronological age by subtracting the child's date of birth from the date of assessment. In doing so, remember:

1. When borrowing days from months, always borrow 30 days regardless of the month.
2. When borrowing months from years, always borrow 12 months.
3. Do not round days of age upward or downward to the nearest month. (A child who is 4 years 11 months and 29 days is still considered a 4 year 11 month old when using the score tables.

4. If a child requires multiple testing sessions, use the first testing date for the age calculation.

To illustrate, the chronological age for a child tested on June 18, 2000 whose birth date is June 20, 1994 is 5 years 11 months and 28 days (see Figure 2.1). You would compare the child's probe scores to the criterion score tables for children 5 years 6 months to 5 years 11 months. Do not round the child's age to 6 years 0 months.

	Year	Month	Day
Date of Assessment	~~00~~ 99	~~6~~ 17	~~18~~ 4
Date of Birth	94	6	20
Chronological Age	5	11	28

Figure 2.1 Calculating Chronological Age

Administration Procedures

General Directions

The following administration, recording, and scoring directions apply to all of the probes.

- Administer the probes in the order of appearance on the *Record Form*.
- Administer the Grammaticality Judgment Probe only to children ages 4.00–8.11. This is a supplemental probe; it is not required to obtain the Elicited Grammar Composite. However, this probe is not appropriate for children younger than age 4.00.
- Administer each item in each probe. There are no basals, ceilings, or other discontinue rules.
- Record each response verbatim, unless otherwise specified. You may tape record the session to facilitate recording responses.
- The elicitation procedures used in this test differ from procedures used in most other language tests. Although the differences may be slight, following these directions exactly is crucial to obtaining accurate results. Review all of the administration directions, especially the information regarding how and when to use the prompts. The Training Videotape will provide you with specific information about the prompts.
- As a general rule for all probes, if you cannot hear or understand what a child said, ask the child to repeat his or her response by saying something like, "Say that again," or "I didn't hear you. Could you repeat it?" Also, you may repeat an item for the child if he or she requests a repetition or if you feel the child was not attending.

Complete administration directions are provided in this manual. Abbreviated directions appear on the Record Form. Once you have carefully read the complete directions and have administered the instrument a few times, you should be able to administer the instrument using the abbreviated directions that appear on the Record Form.

Practice and Trial Items

Practice or trial items are provided for each probe to ensure that the child understands the task before he or she is asked to complete test items. You must present all practice or trial items to the child before administering the test items. If the child does not provide the correct response to a practice or trial item, provide the correct response for the child before continuing with the test items. Follow the directions listed on the Record Form for presenting the practice or trial items for each probe.

Using Prompts

Specific prompts have been developed and are provided for use with each probe to ensure that the correct "response type" is elicited from the child. It is crucial that these prompts be understood and used consistently and correctly during each administration of the Rice/Wexler. The prompts should not be used to encourage a child to change an incorrect response in cases where, although incorrect, the response does represent the type of structure being targeted by the probe. Rather, additional prompting is appropriate to focus the child to attempt a response that includes the structure being targeted. Table 2.1 provides examples of the types of responses that should be probed with additional prompts and the types of responses that should not be followed by additional prompts.

Table 2.1 Examples of Responses to Reprompt and to Not Reprompt

Third Person Singular Probe

Do Reprompt	Do Not Reprompt
He could fly.	He help your teeth.
She will dance.	She dance all around.
He is helping the girl.	He play.
She put the paint on the house.	
. . . helps you.	
. . . play	
He does play.	
He does baseball.	

Past Tense Probe

Do Reprompt	Do Not Reprompt
He was building.	He build it.
She was kicking.	She kick it.
He did it.	
He already finished.	
She is done.	

Phonological Probe

The Phonological Probe is used to determine if the child can produce (or at least mark) the phonemes /*s*/, /*z*/, /*t*/, and /*d*/ in the final position. This is a picture elicitation task. There are 20 items in this probe, which takes about 3 minutes to administer.

Note: "Screening" of the phonemes related to marking grammar prior to evaluating grammar skills is a unique feature of the Rice/Wexler, and it is important that you understand the rationale for the probe and the implications of using this probe correctly before using this instrument. You must administer this probe before administering any Rice/Wexler probe.

The probe is designed to ensure that the child is able to produce the phonemes required to mark grammar. It is not designed to determine articulatory precision or to determine whether or not the child knows the vocabulary associated with the picture; therefore, the administration and scoring for this probe differs from traditional language tasks. For this probe only, you must model the target word for the child if he or she does not use the target word to identify the picture (e.g., rat for "mouse," toothpaste for "squeeze"). Also, provide a model if the child initially does not respond to the item. For you to score the item, the child must attempt the target word. The child is not penalized for receiving a model.

A child's inability to obtain a minimum score of four for each phoneme group tested on this probe indicates that he or she is not able to consistently demonstrate use of the phonemes necessary to mark grammar. It will be impossible in such a case for you to determine whether or not the child's performance on the subsequent probes is a result of a grammar deficit, of phonological involvement, or a combination of both.

Consequently, the Rice/Wexler may not be an appropriate instrument to use with that child and you may not be able to use the scores as described in this manual. While you may want to administer the Rice/Wexler to obtain anecdotal information about the child's language or interactive skills, in this type of case you should not use the results from the Rice/Wexler to determine whether or not the child has a language deficit. Use caution if you choose to proceed with test administration in this situation.

Directions

Say, "I am going to show you some pictures and ask you to name some things." Show the child each picture and ask, "What is this?" or present the prompt below the item on the *Record Form* to elicit the response. If the child does not know the target word, model the word and ask the child to repeat it.

Scoring Responses and Marking the *Record Form*

Read this section carefully before scoring the child's responses. Score only the final phoneme; errors made on other phonemes within the target word do not affect the child's score for this probe.

Score each response as Correct (1 point), Incorrect (0 points), or No Response (NR).

For each item:

Circle 1 if the child marks the final phoneme position spontaneously or in response to your model (either by producing the phoneme correctly, producing a distortion of the target phoneme, or by substituting another phoneme for the target phoneme, such as θ/s, d/t, etc.). Use your clinical judgment to determine if a distortion or substitution indicates that the child is marking the final phoneme.

Circle 0 if the child omits the final phoneme.

Circle NR if the child does not respond.

Note: If the child has several No Response (NR) scores on this section, you may have difficulty eliciting responses throughout the entire Rice/Wexler, which will severely limit the amount of information you will be able to obtain from the instrument.

Completing the Phonological Probe Summary

A. Add the number of correct responses (scores of 1) for each of the four phoneme groups tested and write the child's score (total correct responses) in the appropriate box labeled **Total Score Final / /**.

B. Transfer each Total Score Final / / to the appropriate corresponding box in the Phonological Probe Summary on the bottom of page 2 of the *Record Form*.

C. Place a checkmark in the appropriate box to indicate if the child passed or failed each phoneme group. *Failure in ANY phoneme group results in failure of the entire probe.*

D. Go to the Phonological Probe Result and place a checkmark in the appropriate box (Pass or Fail) to indicate the child's overall performance on the probe.

E. Turn to the Phonological Probe area of the Summary Scores section, located on the front of the *Record Form*. Record the probe result again by placing a checkmark in the appropriate box (Pass or Fail).

Reminder: If a child fails the Phonological Probe, the *Rice/Wexler Test of Early Grammatical Impairment* may not be an appropriate instrument to use to evaluate that child.

Third Person Singular Probe

The Third Person Singular Probe is used to evaluate a child's use of /-s/ or /-z/ in present tense verb forms with singular subjects. This is a picture elicitation task. There is one practice item and 10 test items in this probe, which takes about 5 minutes to administer.

Directions

Say, **"I am going to show you some pictures and ask you to tell me what each person does. Let's try one."** (*Show the picture of the teacher.*) **"Here is a teacher. Tell me what a teacher does."**

Present each item using the standard prompt provided on the *Record Form*. Then for each item, if the child's response

- does not include a subject (e.g., "Make children get well" or "Helps you feel better"), prompt the child further by saying, **"Say a whole sentence," OR "Start with he or she."**
- still does not include a subject after you give the standard prompt, use the alternate prompt: Say, **"Here is a** (occupation)**. Tell me what a** (occupation) **does. A** (occupation) **. . ."** Provide this alternate prompt only once for each item. Wait for the child to complete the response, then record the response including the subject you provided.
- includes a plural subject such as "They put out fires," prompt further by saying, **"Tell me just what this** (occupation) **does,"** OR **"Start with he or she."**
- is ambiguous or uses a different structure or verb tense such as "He is working," say, **"Yes, he is working, but tell me what he does."** Subsequently, you may also follow up with **"Start with he or she."** Provide this prompt only once per item. If the child still does not provide the targeted form, record the response verbatim and proceed to the next item.

- includes a subject and the targeted third person singular structure, but is incorrect, *do not* present additional prompts. Record the response verbatim and proceed to the next item. For example, if the child says, "A dentist *fix* your teeth" instead of "A dentist *fixes* your teeth," record this complete response and proceed to the next item.

Note: Some children prefer to use "it" as the subject. If this occurs, on the first use of "it" as the subject, say to the child, **"Start with he or she."** If the child persists with using "it," accept this as the subject and proceed with the other items. Responses that use "it" as the subject and include the third person singular -s ending are scored as correct.

Recording Responses

Record the child's entire response *verbatim* for each item. Listen carefully as the child responds and make sure that what you record reflects the exact response given. You may record responses phonetically or orthographically when completing the *Record Form*. To assist you in honing your listening skills as you administer this probe, the following are examples of children's (with and without impairment) responses to these items.

Teacher—A teacher write on board.
A teacher teaches.

Dentist—He fix your teeth.
It checks your teeth and heart.

Dancer—She dances and twirls.

Scoring Responses and Marking the *Record Form*

Mark each response as Correct, Incorrect, Unscorable, or No Response, by placing a checkmark in the appropriate column.

Note: It is not necessary that the content of the response be accurate, nor is it necessary that the child's response relate to the stimulus picture. Responses such as "A dentist helps your head" or "A dentist goes to the moon" are considered correct. You would place a checkmark in the Correct column for either of these responses.

For each item, place a checkmark in the:

Correct column for a response that includes a third person singular subject (provided by the child or by you) and a correct third person present tense singular verb form (e.g., He fixes, The painter paints, It twirls around and dances).

Incorrect column for a response that includes, or appears to include, an attempt of a third person singular verb form (e.g., He spray, She fly), but omits the -s **OR** includes a double marking of the verb (e.g., He testses, She playses, He foughts, She throwses, He giveses).

Unscorable column for a response that includes any verb form or tense other than the third person singular present tense, whether the other verb form is correct or incorrect (e.g., She played, He will help, She is working, They help, She does help, He does fix).

No Response column for an item to which the child does not respond.

Refer to Appendix E for additional examples of scoring Third Person Singular items.

Completing the Third Person Singular Probe Summary

A. Count the number of checkmarks in each column and record the number in the Total Score box at the bottom of the appropriate column, following these steps:
   - record the total of checkmarks in the Correct column in the Total Score box labeled **A**, and then
   - record the total of checkmarks in the Incorrect column in the Total Score box labeled **B**; next
   - record the total of checkmarks in the Unscorable column in the Total Score box labeled **U**, and finally
   - record the total of checkmarks in the No Response column in the Total Score box labeled **NR**.

B. Transfer the total of the Correct column (A) to the Third Person Singular Probe Summary box labeled **A**. This is the numerator for calculating the probe score.

C. Add the totals of score box A and score box B. Write this value in the Third Person Singular Probe Summary box labeled **Sum of A + B**. This is the denominator for calculating the probe score.

D. Calculate the Third Person Singular Probe score by dividing the value of the numerator by value of the denominator. (You may also refer to Appendix C and obtain the score by finding the number at the intersection of these two values.) Write this number in the box labeled **Third Person Singular Probe Score**.

See Figure 2.2 for an example of recording and scoring responses for the Third Person Singular Probe.

Past Tense Probe

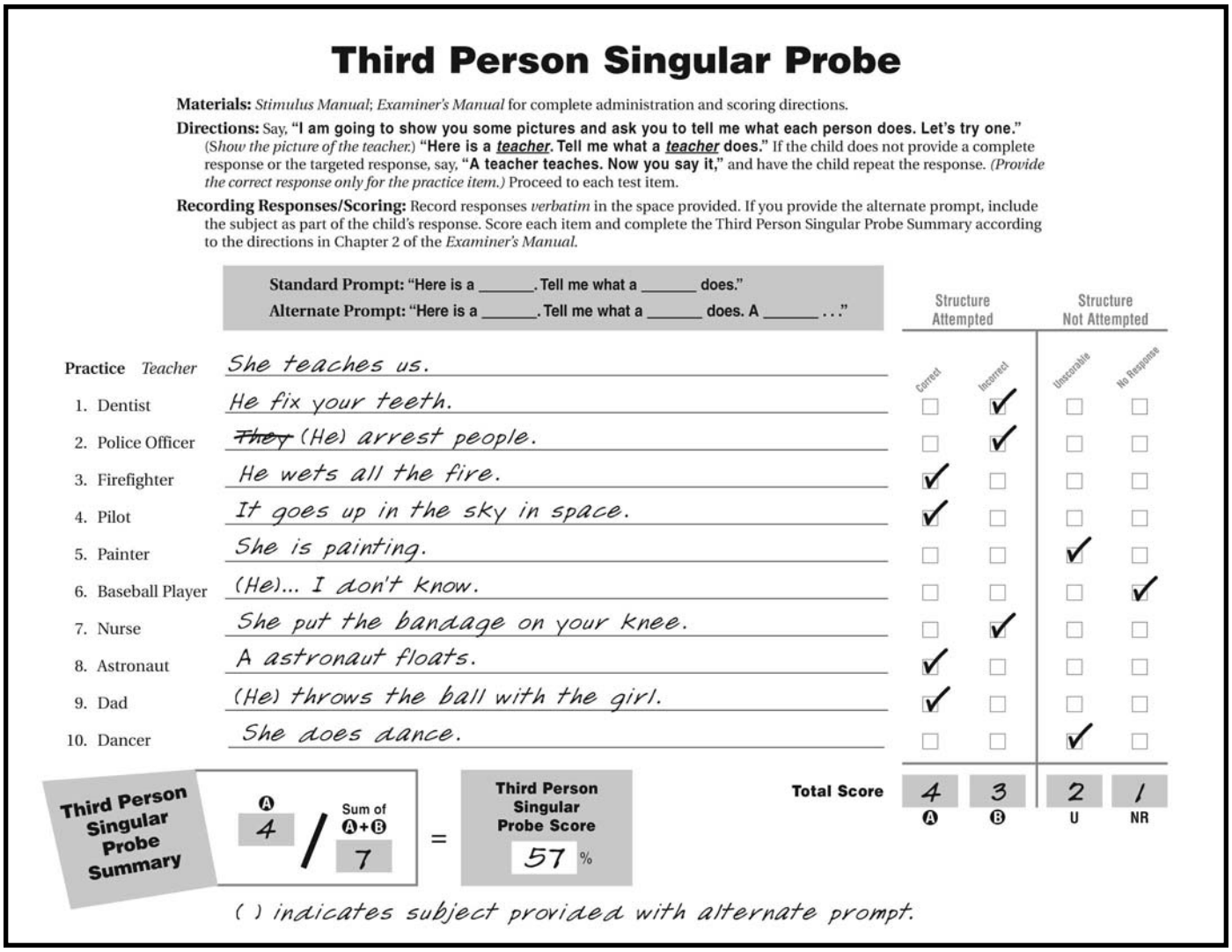
The Past Tense Probe is used to evaluate a child's use of regular past tense (-ed) verb forms and irregular past tense verb forms. This is a picture elicitation task. There are two practice items and 18 test items in this probe, which takes about 5 minutes to administer.

Directions

Say, **"I have two pictures. I will describe the first one and you tell me about the second one. Let's try one."** (*Point to raking picture.*) **"Here the boy is raking.** (*Point to raked picture.*) **Now he is done. Tell me what he did."**

Present each item to the child using the standard prompt provided on the *Record Form*. Then for each item, if the child's response

- does not include a subject (e.g., "Raked the leaves"), prompt further by saying, **"Say a whole sentence," OR "Start with he or she."**
- still does not include a subject after you give the prompt, use the alternate prompt: Say, **"Here the boy/girl is _____. Now he/she is done. Tell me what he/she did. He/She . . ."** Provide this alternate prompt only once for each item. Wait for the child to complete the response, then record the response including the subject you provided.



Third Person Singular Probe

Materials: *Stimulus Manual*; *Examiner's Manual* for complete administration and scoring directions.

Directions: Say, **"I am going to show you some pictures and ask you to tell me what each person does. Let's try one."** (*Show the picture of the teacher.*) **"Here is a *teacher*. Tell me what a *teacher* does."** If the child does not provide a complete response or the targeted response, say, **"A teacher teaches. Now you say it,"** and have the child repeat the response. *(Provide the correct response only for the practice item.)* Proceed to each test item.

Recording Responses/Scoring: Record responses *verbatim* in the space provided. If you provide the alternate prompt, include the subject as part of the child's response. Score each item and complete the Third Person Singular Probe Summary according to the directions in Chapter 2 of the *Examiner's Manual*.

Standard Prompt: "Here is a _____. Tell me what a _____ does."
Alternate Prompt: "Here is a _____. Tell me what a _____ does. A _____ . . ."

		Structure Attempted		Structure Not Attempted	
		Correct	Incorrect	Unscorable	No Response
Practice *Teacher*	She teaches us.				
1. Dentist	He fix your teeth.	☐	✓	☐	☐
2. Police Officer	~~They~~ (He) arrest people.	☐	✓	☐	☐
3. Firefighter	He wets all the fire.	✓	☐	☐	☐
4. Pilot	It goes up in the sky in space.	✓	☐	☐	☐
5. Painter	She is painting.	☐	☐	✓	☐
6. Baseball Player	(He)... I don't know.	☐	☐	☐	✓
7. Nurse	She put the bandage on your knee.	☐	✓	☐	☐
8. Astronaut	A astronaut floats.	✓	☐	☐	☐
9. Dad	(He) throws the ball with the girl.	✓	☐	☐	☐
10. Dancer	She does dance.	☐	☐	✓	☐
Total Score		4 (A)	3 (B)	2 (U)	1 (NR)

Third Person Singular Probe Summary: Ⓐ 4 / Sum of Ⓐ+Ⓑ 7 = **Third Person Singular Probe Score** 57 %

( ) indicates subject provided with alternate prompt.

Figure 2.2 Recording and Scoring Responses for the Third Person Singular Probe

- is ambiguous or does not include a past tense structure, such as, "He's done" or "He finished,"* say, **"Yes, he's done (or he finished), but tell me what he did to the leaves, or use the same word I use."** Provide this prompt only once per item. If the child still does not provide the targeted form, record the response verbatim and proceed to the next item.

**Note: The response "He finished," although technically a correct regular past tense construction, is sometimes used repeatedly by some children. In this case, you would reprompt and attempt to have the child use a different verb so that you can better evaluate his or her ability to construct a regular past tense. If a child uses this response repeatedly, after three uses of this response, count additional uses as Unscorable. This rule was used in the standardization research and prevents a child from obtaining an inflated score on the basis of repeating one response multiple times. You should also use this rule if the child repeats any other response (whether correct or incorrect) more than three times.*

Because we are interested in the child's ability to use a past tense structure, rather than in testing a specific verb, it is not necessary that the child use the same verb shown in the stimulus picture; it is important, however, that the child use a verb that is the same type of verb (i.e., regular or irregular) as the target. If the child provides a past tense response with a different verb of the same type (e.g., *washed* instead of *cleaned, built* instead of *made, hop* instead of *jumped*—the latter is an example of a failed attempt at the same-type verb, hopped), then regardless of whether or not the verb is inflected properly, do not provide additional prompts. Record the child's response and proceed to the next item.

However, if the child substitutes an irregular verb for a regular verb or vice versa, then regardless of whether or not the verb is inflected correctly, (e.g., "She spelled" instead of "She wrote"), prompt the child once with "Use the same word I use." Then readminister the item, record the child's response, and proceed to the next item.

Although *ideally* the child should provide the type of response being elicited, some children will not be able to do this. Ultimately, you will score whichever response type the child gives after appropriate prompts, even if it does not match the target.

Recording Responses

Record the child's entire response *verbatim* for each item. Listen carefully as the child responds and make sure that what you record reflects the exact response given. You may record responses phonetically or orthographically when completing the *Record Form*. To assist you in honing your listening skills as you administer this probe, the following are examples of children's (with and without impairment) responses to these items.

caught—She caught the ball.
She catched the ball.

made—He made the birdhouse.
He make the birdhouse for the bird.

climbed—She climbed.
She climbeded.

rode—He rided.
He roded it.

dug—She dug.
She digged a big hole.

ate—He ate.
He eated.

Scoring Responses and Marking the *Record Form*

Score regular verbs separately from irregular verbs, as shown on the *Record Form*. To assist you with this distinction, irregular verbs appear on the *Record Form* in italics. Mark each response as Correct, Incorrect, Unscorable, or No Response by placing a checkmark in the appropriate column.

Note: *It is not necessary that the content of the response be accurate, nor is it necessary that the child's response relate to the stimulus picture. To illustrate, for responses to the first item of this probe (painted) such as, "The boy nailed the fence" or "The boy washed the house," you would enter a checkmark in the Correct column for Regular Verbs. If a child responds with "The boy built the house" (i.e., with an irregular verb rather than the*

targeted regular verb) reprompt for a regular verb (the targeted verb type). If the child persists with an irregular verb response you will place a checkmark in the Correct column for Irregular Verbs, and proceed to the next item.

For each response that includes a Regular Verb, place a checkmark in the:

Correct column for a response that includes a subject (provided by the child or by you) and any correctly formulated regular past tense verb (e.g., He painted; The girl cleaned; He brushed).

Incorrect column for a response that includes, or appears to include, an attempt of any regular past tense verb, but does not correctly formulate it or omits the -ed (e.g., He paint, The girl clean, He brush).

For each response that includes an Irregular Verb, place a checkmark in the:

Correct column for each response that includes a subject (provided by the child or by you) and any correctly formulated irregular past tense verb form (e.g., He made; She wrote; He gave).

Overregularization (Overreg.) column for each response that includes a subject (provided by the child or by you) and an overregularization. An overregularization is an irregular past tense verb that includes the irregular stem with a regular past affix (e.g., gived, maked, catched, etc.) or a correct irregular past tense construction with a regular past affix (e.g., wroted, maked, gaved).

Incorrect column for each response that includes an attempt of an irregular verb that is not formulated correctly (e.g., write, give, ride).

If a Past Tense Structure was not attempted, place a checkmark in the:

Unscorable column for a response that includes any other verb tense (other than past), whether it is correct or incorrect (e.g., I can do that, He will paint the house, He is done).

No Response column for an item to which the child does not respond.

Refer to Appendix E for additional examples of scoring Past Tense Probe items.

Completing the Past Tense Probe Summary

A. Count the number of checkmarks in each column and record the number in the Total Score box at the bottom of the appropriate column, following these steps:

- record the total of checkmarks in the Regular Verbs Correct column in the Total Score box labeled **A**, and then
- record the total of checkmarks for the Regular Verbs Incorrect column in the Total Score box labeled **B**; next,
- record the total of checkmarks in the Irregular Verbs Correct column in the Total Score box labeled **C**,
- record the total of checkmarks in the Irregular Verbs Overregularization column in the Total Score box labeled **D**, and then
- record the total of checkmarks for the Irregular Verbs Incorrect column in the Total Score box labeled **E**; next,

- record the total of checkmarks in the Unscorable column in the Total Score box labeled **U**, and finally,
- record the total of checkmarks in the No Response column in the Total Score box labeled **NR**.

B. Sum the totals of columns A, C, and D (Regular Verbs Correct, Irregular Verbs Correct, and Irregular Verbs Overregularization). Transfer this number to the box in the Past Tense Probe Summary labeled **Sum of A + C + D**. This is the numerator for calculating the probe score.

C. Sum the totals of columns A, B, C, D, and E (Regular Verbs Correct, Regular Verbs Incorrect, Irregular Verbs Correct, Irregular Verbs Overregularization, and Irregular Verbs Incorrect). Transfer this number to the box in the Past Tense Probe Summary labeled **Sum of A + B + C + D + E**. This is the denominator for calculating the probe score.

D. Calculate the Past Tense Probe score by dividing the value of the numerator by the value of the denominator. (You may also refer to Appendix C and obtain the score by finding the number at the intersection of these two values.) Write this number in the box labeled **Past Tense Probe Score**.

Supplemental Scoring for the Past Tense Probe

In addition to the above scoring that is required to calculate the Past Tense Probe score for the Elicited Grammar Composite, you may also want to examine the specific types of past tense responses (regular past tense, irregular past tense, and irregular past finite) the child uses. A child's use of these three structures provides valuable additional information that can be used for diagnosis and intervention planning. See Chapter 3 for interpretive information regarding these scores.

To examine a child's performance on specific past tense structures, you will use the Supplemental Scoring section of the Past Tense Summary area in the *Record Form*. To use this section, transfer the Total Scores for Regular Verbs Correct (A), Regular Verbs Incorrect (B), Irregular Verbs Correct (C), Irregular Verbs Overregularization (D), and Irregular Verbs Incorrect (E) to the supplemental scoring area, as indicated, to calculate the Regular Past score, the Irregular Past score, and the Irregular Past Finite score. Calculate each of these scores above by dividing the numerator by the denominator. You may also refer to Appendix C to obtain the score by finding the number at the intersection of these two values. Then refer to the interpretation section of Chapter 3 for direction regarding how to use this information in the clinical process.

See Figure 2.6 for an example of recording and scoring responses for the Grammaticality Judgment Probe.

Completing the Summary Scores

Calculating the Screening Score

A. Transfer the probe scores for the Third Person Singular Probe and the Past Tense Probe to the Screening test area on the Summary Scores section on page 1 of the *Record Form*.

B. Sum the scores for these two probes.

C. Record this score in the box labeled **Sum of Screening Probe Scores**.

Past Tense Probe

Materials: *Stimulus Manual*; *Examiner's Manual* for complete administration and scoring directions.

Directions: Say, **"I have two pictures. I will describe the first one and you tell me about the second one. Let's try one."** *(Point to raking picture.)* **"Here the boy is raking.** *(Point to raked picture.)* **Now he is done. Tell me what he did."** Wait for the child's response.*

Then say, **"Let's try another one."** *(Point to skating picture.)* **"Here the girl is skating. Now she is done.** *(Point to skated picture.)* **Tell me what she did."** Wait for child's response then proceed to each test item.*

*If the child does not respond correctly to either practice item, complete the item(s) for the child by saying, **"He raked/She skated."** *(Provide the correct responses only for practice items.)*

Recording Responses/Scoring: Record responses *verbatim* in the space provided. If you provide the alternate prompt, include the subject that you provided as part of the child's response. Score each item and complete the Past Tense Probe Summary according to the directions in Chapter 2 of the *Examiner's Manual*.

Standard Prompt: "Here the boy/girl is _____. Now he/she is done. Tell me what he/she did."
Alternate Prompt: "Here the boy/girl is _____. Now he/she is done. Tell me what he/she did. He/She . . ."

		Structure Attempted: Regular Verbs: Correct	Structure Attempted: Regular Verbs: Incorrect	Structure Attempted: Irregular Verbs: Correct	Structure Attempted: Irregular Verbs: Overregularization	Structure Attempted: Irregular Verbs: Incorrect	Structure Not Attempted: Unscorable	Structure Not Attempted: No Response
Practice 1. raked	He is raking. (prompt) He raked.							
Practice 2. skated	She skated.							
1. painted	He painted the fence.	✓						
2. *caught*	She caught the ball.			✓				
3. *made*	He built the birdhouse.			✓				
4. brushed	He combed his hair.	✓						
15. *blew*	She blow the candle.					✓		
16. tied	The girl ties her shoe.						✓	
17. lifted	Her lifted the box.	✓						
18. *gave*	He gave his mom a present.			✓				
Total Score		7 (A)	1 (B)	3 (C)	3 (D)	1 (E)	2 (U)	1 (NR)

Past Tense Probe Summary

Sum of (A+C+D): 13 / Sum of (A+B+C+D+E): 15 = **Past Tense Probe Score**: 87 %

Past Tense Supplemental Scoring

Regular Past Tense	A: 7 / Sum of (A+B): 8	=	88 %
Irregular Past Tense	C: 3 / Sum of (C+D+E): 7	=	43 %
Irregular Past Finite	Sum of (C+D): 6 / Sum of (C+D+E): 7	=	86 %

Figure 2.3 Recording and Scoring Responses for the Past Tense Probe

D. Calculate the Screening Test score by dividing the Sum of Screening Probe Scores by two. Record this score in the box labeled **Screening Test Score**.

Identifying Criterion Scores

Criterion scores are provided in Appendix A for each probe score (including the scores for the supplemental probe—Grammaticality Judgment) and for the Elicited Grammar Composite. Criterion scores are organized by each half-year age interval from age 3.00–8.11.

Identify the criterion score for each probe score. Locate the appropriate criterion score table for each probe for the child's age. Record the criterion score for each probe on page one of the *Record Form*, in the column labeled **Criterion Score**. Then, place a checkmark in the appropriate column next to the criterion score to indicate that the child's score is either "At/Above Criterion" or "Below Criterion."

Marking the Growth Curves

Growth curves are provided to enable you to determine how the child's performance on the Elicited Grammar Composite and on each individual score of the Rice/Wexler compares against his or her age-level peers. The growth curve for the Elicited Grammar Composite appears on the front of the *Record Form*. Growth curves for individual probe scores appear on pages 18 and 19 of the *Record Form*.

Use of growth curves is optional. To use the growth curves, plot the child's probe score (not criterion score) onto the graph. Find the child's age group (ages are shown in half-year intervals) across the bottom of the curve. Then locate the child's probe or Elicited Grammar Score on the vertical axis in the graph. Mark the graph at the intersection of these two values/points.

To use the Growth Curves most effectively it is important to understand that the criterion scores represent the point at which there is good sensitivity and as well as good specificity. In other words, there is a good chance of identifying a child with a language disorder while at the same time being able to identify those children who do not have a language disorder. This latter criterion ensures that the criterion scores will be at lower percentiles, i.e., where most of the normal children scored above that level. When looking at the graphs this may result in a plot that, although a child scores above the criterion, when his or her scores are plotted they fall below (sometimes well below) the percentiles plotted on the graph. This is to be expected, given the nature of the development of the criterion scores. The percentiles, then, serve to give you an idea of how the children in the normal group performed, but is not an indicant that children must perform in this range to be considered normal.

Also note that the growth curves show performance of the children in the normal language group. For this study, the upper age of the children in this group was 6.11. Therefore there are no comparisons available for children 7 and 8 years of age in the normal group. Use your clinical judgment to determine if comparing the performance of a 7- or 8-year-old to a younger normal group would be beneficial.

Refer to the information provided in Chapter 3 for guidelines for interpreting each of these scores.

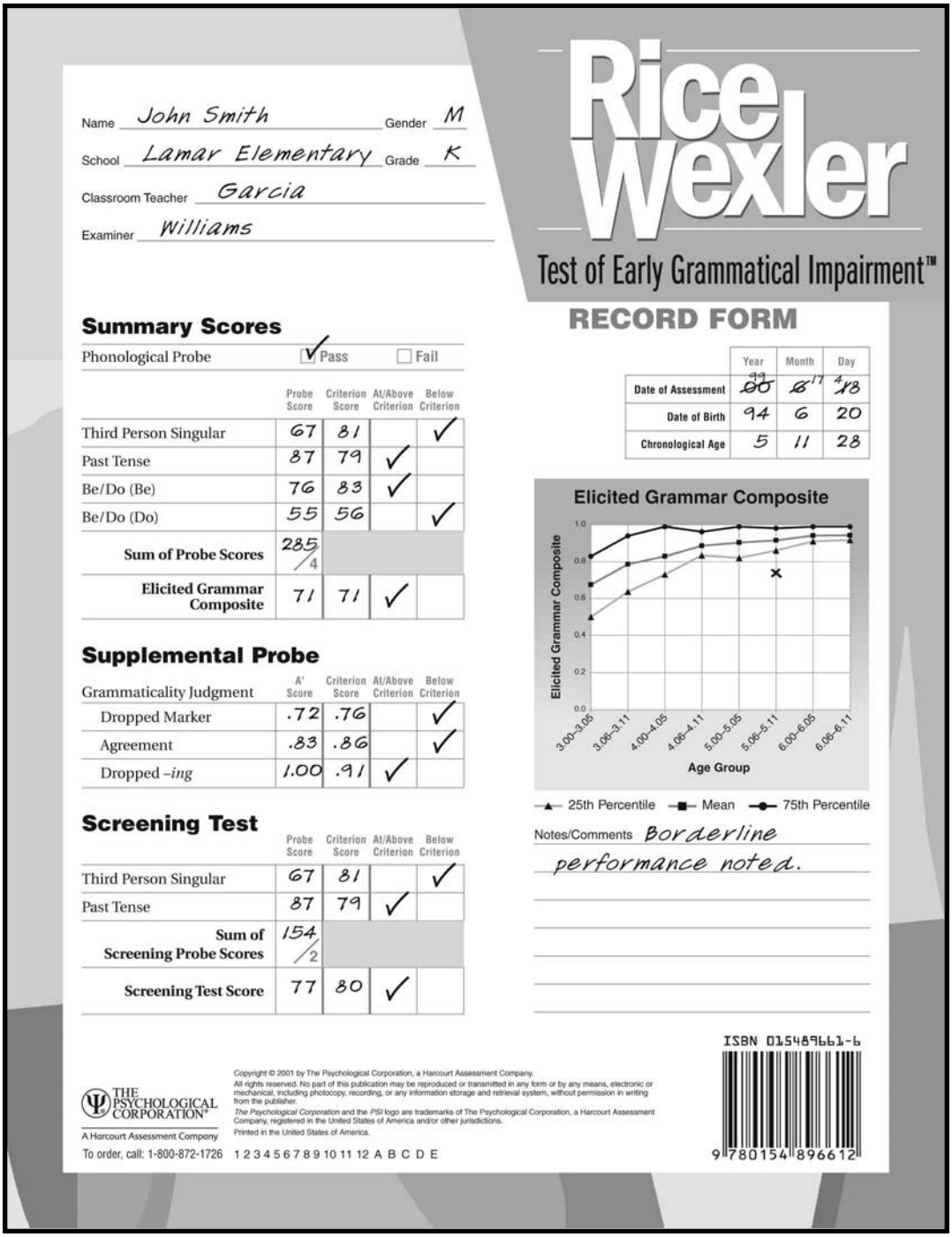


Figure 2.7 Completing the Summary Scores Sections

Interpretation 3

The *Rice/Wexler Test of Early Grammatical Impairment* offers a wealth of information to describe a child's morphosyntactic competencies and limitations; to determine if a child's knowledge is within the expected, age-appropriate range; to identify target areas for intervention; and to evaluate change over time.

Scores Calculated

Seven individual scores can be calculated for the R*ice/Wexler Test of Early Grammatical Impairment*. Four of these scores contribute to the Elicited Grammar Composite score. These scores include the Third Person Singular Probe score, the Past Tense Probe score, the Be score, and the Do score. The three remaining scores are generated from the Grammaticality Judgment Probe, which is a supplemental probe. These scores include the Dropped Marker score, the Agreement Score, and the Dropped *-ing* score.

Supplemental Scoring can also be calculated for the Past Tense Probe and for the Be/Do Probe. This level of scoring is not required to obtain the Rice/Wexler test results, but is provided for those users who wish to obtain more in-depth information about a child's performance. Using the supplemental scoring, you can examine a child's performance on specific structures. This information can be helpful in the diagnostic process and for planning intervention.

The Rice/Wexler can be used also as a screening tool. To obtain a Screening Test score, follow the directions provided in Chapter 2 of this manual.

In addition to the probes scores, criterion scores can be obtained for each probe score, for the Elicited Grammar Composite, and for the Screening Test. Finally, growth curves are provided to enable you to interpret a child's score on each probe relative to children of the same age.

About the Scores

Each score obtained for the Rice/Wexler represents a percentage correct. In other words, the scores represent the number of items answered correctly relative to the number of items attempted by the child. This contrasts with most available instruments where the number attempted is the number of items that were administered. This usually includes items the child did not respond to and any off-topic responses. For the Rice/Wexler, however, the number of items attempted includes only those items in which the child actually attempted the particular structure being elicited. Items that the child does not respond to, or that the child provides a structure type different from what was being elicited, do not enter into the score calculations.

Rice/Wexler scoring is designed to provide a precise estimate of what a child knows about obligatory properties of a grammatical marker, independent of that child's choice of lexical item, off-task behavior, or understanding of non-grammatical elements of a particular item. For example, if we are interested in a child's knowledge of the need to provide a past tense form, the fact that the child may sometimes fail to respond or may give an off-topic answer (responses referred to as "unscorable") is irrelevant to the estimate and, for that reason, is not included in the scoring. The many ways that a child can fail to give the target response are included in a different measure, one that produces a total of unscorable items. This measure can provide insights into a child's level of accuracy, given the full range of processing demands, off-task behaviors, and other reasons for low performance, in addition to his or her lack of understanding regarding the grammatical structures assessed by the test.

It is also important, under these conditions, to be aware of the number of attempts a child makes on a particular probe. The interpretation of the performance level for a child who attempts only one or two structures may be very different from a child who attempts ten structures, even though both children may receive the same score. In general, fewer attempts and more off-task responses are typical of children with more immature language; this may be because a child is very young or because a child's language level is not commensurate with the demands of the task. In these cases, a record of the number of attempts can be a useful adjunct to a child's score on the probe because as a child's language competence grows the number of attempts also tends to increase.

Development of Criterion Scores

Criterion scores are provided for each probe. These criterion scores provide or represent "cut points" and enable you to determine or report whether or not a child has performed within the normal range, as indicated by the test results. The rationale used to determine the cut points involved consideration of the bimodal distribution of affectedness, as illustrated in Figure 1.2 in Chapter 1. Selection of the criterion scores involved considering for each age level where the tails of each distribution overlap between the language disorder group (on the low end) and the normal group (on the high end). The cut point represents the performance level that best separates the two groups.

The criterion scores were developed by the authors, utilizing clinical experience and expertise, and informed by the data collected for this test. As a general guideline, the selection of criterion scores was based on providing at least 80% sensitivity, following the belief that it is best to try to identify children who score at or below 80% of the clinical group, for a given age level. The corresponding specificity and the relationships of criterion scores across age were also taken into account, meaning that fluctuations in the data due to sampling error were considered when developing the scores. Developing cut points in this way yields points that follow a general upward progression as the children get older and that generally have a specificity of .80 or higher. As expected, specificity also increases with age. The criterion scores provided here represent recommendations for appropriate cut points for each age level. For those users who wish to use or develop alternate cut points or criterion scores that may be more appropriate for a particular situation, tables of obtained sensitivity and specificity for each possible probe score (0–100) are also provided (see Appendix B).

Although criterion scores can be obtained for each probe, three criterion scores can be of particular clinical relevance. The criterion score for the screening test is important for quickly determining whether or not a child needs further evaluation. For children

who receive the complete Rice/Wexler, the Elicited Grammar Composite provides a summary score for the four individual probes that contribute to it, i.e., Third Person Singular, Past Tense, Be, and Do. This score can be used in conjunction with results from a complete assessment battery to determine if a diagnosis of language impairment is appropriate for a child. Of the Grammaticality Judgment indices, the A' score for Dropped Marker would be considered a corollary to the Elicited Grammar Composite production index (see discussion below).

Using the Rice/Wexler in Clinical Practice

The Rice/Wexler is appropriate for use with a wide range of children. This includes children who seem to be developing normally and may have good speech skills, but for whom there is a question about whether or not their level of language development is appropriate for their age; children who have been regarded as "socially immature" or "slightly AD/HD," and who may have an underlying, undetected language impairment that inhibits their social growth; children at the age of kindergarten entry, for whom there are questions of their readiness to attend school; children who have been identified as developmentally delayed, but there is little specific information about their grammatical abilities; and children with a diagnosis of "mild autism," who may have special grammatical deficits as part of a language impairment.

The intended use of the Rice/Wexler is to provide a detailed description of a child's performance in finiteness marking, which is an area of morphosyntax that differentiates children with language impairments from children without language impairments. Because the Rice/Wexler does not examine all elements of a child's language growth and competency, it is important that a child receive a full assessment that includes other tests, instruments, or procedures that provide the information necessary for a full diagnosis of language impairment. Assessments of receptive and expressive vocabulary and assessments of other elements of grammatical development should be included in such an assessment. In addition, spontaneous language sampling would provide important information about a child's mean length of utterance; tasks that provide information about discourse or narrative skills would also be appropriate. This full array of information would enable you to determine whether or not the results of the Rice/Wexler are part of more pervasive language limitations for a child, or if a child's performance on other indices, the lexicon for example, is more commensurate with age expectations.

Performance on the Rice/Wexler can be interpreted in a relatively straightforward manner as how much a child knows about the need to mark finiteness in simple clauses. Other language tests may focus on a more general sample of grammatical skills or may focus heavily on semantics or vocabulary development. Because the property of finiteness in affected children is not necessarily tightly linked to other language properties (it can be relatively weak or strong), it is possible for a child to have a relatively low performance on the Rice/Wexler, while at the same time show relative strengths on other language tests. The exact interpretation of the differing results depends on the particular language tests used and on how easily one can determine the nature of the language skills being measured. It may be necessary to carefully look at individual items or subtests to determine the source of the differences. Keep in mind also that the Rice/Wexler is designed to estimate the likelihood of finiteness marking, while other tests may include only one or two items that perform this function. For this reason, a child who sometimes uses a morpheme and sometimes does not may get either too much credit on a conventional test (if the occasional usage appears in the test session) or not enough credit (if the occasional non-usage appears in the test session).

Establishing Eligibility for Services

Eligibility for services is often defined in terms of a child's performance on an omnibus test, with performance criteria such as "one standard deviation below the age mean" or "at the 16th percentile or below." The reference points for these criteria are based on a theoretically normal distribution of non-clinical children of the same age, i.e., the normal group in Figure 1.1 in Chapter 1. The results obtained on the Rice/Wexler enable you to make the same determination for establishing eligibility. The data obtained enable you to describe a child's performance in terms of the normal group (i.e., specificity) or the disorder group (i.e., sensitivity). A specificity criteria of 84% will give a value that 84% of the normal group scored at or above. This value identifies the level at which the bottom 16% of the normal group scored, which is roughly comparable to the 16th percentile values reported on a typical omnibus test. Conversely, a good argument could be made that children who score at or below the level of 84% of the clinical group (i.e., a sensitivity of .84) should be considered eligible for services, especially if this score has a high level of specificity. For example, when using the Rice/Wexler as a screening tool and following the recommended criterion scores for children ages 4.06–4.11, a sensitivity of .86 has a specificity of .94, suggesting that a child who obtains this score is both in the clinical range and not in the normal range of performance.

Using Rice/Wexler as a Screening Tool

Using the Rice/Wexler as a screening tool will enable you to obtain a quick estimate of a child's skills to determine whether or not a more comprehensive assessment is warranted. This evaluation can be completed in about 10 minutes. To use the Rice/Wexler as a screening tool, follow the directions provided in Chapter 2.

The Screening tool is composed of two probes from the complete test: the Third Person Singular Probe and the Past Tense Probe. These probes were selected for use as a screening tool based on their generally high sensitivity and specificity values, their relative ease and quickness of administration, and their familiar picture elicitation format. Criterion scores for the screening test are provided in Appendix A. The criterion scores for the screening test were determined in the same way as the criterion scores for the other probes.

A child who fails the screener should be considered at risk for school readiness and scheduled for a full language assessment to determine if language intervention or specialized services are warranted. For young children whose performance is considered borderline (scores that are close to the criterion score, whether above or below it), you may want to administer the screening again a few months later, to determine if later performance meets age expectations. If it does not, a full assessment should be carried out. If you suspect situational factors (such as fatigue) may have contributed to low performance during the screening, it would be appropriate to re-administer the probes at a later time to determine whether or a not a child performs differently on subsequent testing.

It is expected that using the screening probes will be especially helpful to clinicians, in those regions of the country that carry out school-entry screenings to determine a child's developmental status and readiness for school, in identifying children that are about to enter kindergarten or first grade who may have a language impairment.

Interpreting the Rice/Wexler Scores

Elicited Grammar Composite

The Elicited Grammar Composite represents a child's performance in the four areas of grammar tested by the Rice/Wexler: Third Person Singular, Past Tense, copulas and auxiliaries of Be, and auxiliaries of Do. These four areas together constitute a grammatical marker. The Elicited Grammar Composite can be considered the total test score that indicates the likelihood that a child will supply an obligatory marker of finiteness.

Use the criterion scores to determine if the child's grammatical skills in the area of finiteness marking appear to be on-target or deficient. A child who scores above the criterion score for his or her age appears to be progressing toward adult grammar at the expected rate. This child probably would not need additional evaluation in the area of grammar skills. A child who scores below the criterion score may not be progressing toward the adult grammar in a timely manner. This child may be in need of further evaluation. As noted above, a full language assessment is recommended to determine if the child has a language impairment and, if so, the pervasiveness of that impairment.

Interpreting the Phonological Probe Results

The Phonological Probe provides information about a child's ability to produce the phonemes needed for the morphemes tested in the Rice/Wexler, and is included to screen those children whose phonological impairments may yield ambiguous results on the test.

If a child does not produce a final /z/ for the inflected verb "goes" in the sentence "Patsy goes home," it could be because the child does not know that the /z/ is required to mark present tense with third person singular subjects (i.e., the morphosyntactic knowledge), or it could be because the child does not know how to produce final /z/. It is important not to confuse these two different competency levels, although they are obviously interrelated; one is more likely a morphosyntactic problem and one is more likely a phonological problem.

It is possible for children to know the phonological properties of /z/ but not to know the morphosyntactic requirements for morphemes expressed as /z/, such as in plurals ("bugs") and in the third person singular present tense (goes). The Rice/Wexler focuses on morphosyntactic competency. Within this criterion, it does not matter if a child uses a distorted /s/ or /z/, for example. What is important is that there is a consistent, recognizable sound used in place of /s/ or /z/. It is important that you have this information for each child before you proceed with administration of the test.

If a child passes this probe, you know that the child has the capability to produce the phonemes associated with the morphemes being tested. You may then proceed with administration of the other probes and be reasonably certain that errors demonstrated on this test are indicative of a morphosyntactic problem. If a child does not pass this probe, a more complete assessment of phonological development is indicated. It is important to ascertain whether or not a child's phonological development is sufficient to express morphological distinctions. Depending on the results of this assessment, you may want to proceed only with certain probes of the Rice/Wexler, or you may want to proceed with the complete administration.

It could be, for example, that a child has adequate control of /s/ and /z/, but not /t/ and /d/. In this case, it may be appropriate and useful to administer the Third Person Singular Probe and the Be/Do Probe, where the /s/ and /z/ phonemes are needed. For

the purpose of obtaining some information about past tense, you may also complete the Past Tense Probe, but focus only on the irregular past tense items, where final /t/ and /d/ do not represent past tense, i.e., for items such as "catch," if a child says "caugh" you can still tell that he or she attempted a past tense form of the verb, even though phonological mastery of the final /t/ is not yet complete.

You may also want to administer the Grammaticality Judgment Probe to see if a child's judgments of morphosyntax are commensurate with age expectations, even if phoneme production is inadequate (which suggests that the child's difficulties are more attributable to phoneme production than to morphosyntactic limitations). Obtaining information on certain probes only may be useful for specific clinical purposes, but should not be interpreted as having obtained a complete assessment of the grammatical markers associated with this type of language disorder.

Interpreting the Third Person Singular Probe Results

The Third Person Singular Probe examines the child's ability to produce a third person singular structure. The task is designed to elicit a complete sentence that includes a singular subject and a verb. It is crucial to elicit a complete sentence response to obtain the information necessary to make an accurate judgment about a child's abilities in this area. You need to elicit "He works" or "He work," as opposed to "work," because including the subject makes the addition of the affix -s necessary if the sentence is to be grammatically well-formed. Without a subject, the child could generate a bare-stem verb form , i.e. "work," as a citation verb form that describes the activity (which is allowable in the elliptical response of adult grammar).

The preferred elicitation procedure is designed to elicit a spontaneous response from the child and to avoid providing a model for the child to imitate. The acceptable alternate elicitation procedure involves the use of a prompt in the form of a cloze procedure. In this case, you provide the subject and the child completes the rest of the response. These procedures ensure that the child's response is based on his or her grammatical system or skills and provides you with a clearer view of the child's capabilities and limitations in this area.

As described elsewhere in this manual, only those responses for which the child attempts a complete sentence with an overt subject and a lexical verb (and that attempts a third person singular structure) are included in the Third Person Singular Probe score. Responses that include other structures or verb tenses are unscorable and do not affect the scores positively or negatively; these items are disregarded when determining the probe score. In addition, there are no specific targeted lexical verbs: the child may use any lexical verb that he or she chooses. What matters is that a regular lexical verb that requires the third person singular -s is used. This criterion excludes auxiliary and main verb uses of "has" because "has" is generally considered to be an irregular form (i.e., have/has/had).

Two advantages of the scoring procedures are that a child can get credit for using different verbs and that they provide a context where the necessity of the targeted morpheme to the integrity of the elicited structure is as unambiguous as possible. In other words, the task requires that a child adopt the expected structures in his or her responses. Some children need additional prompts to fully understand the task, and instructions for using these prompts are provided in the administration and scoring section of Chapter 2. However, even with additional prompts, some children may not provide scorable responses. This may indicate that the child is functioning at a younger level of language competency than is expected; that he or she is showing

signs of fatigue, disinterest, or noncompliance with the assessment; or that the child is stuck on a preferred but non-target response, such as "He is working." As with any testing scenario, if you suspect fatigue, disinterest, or noncompliance, you may try re-administering the probe on another day. This may help determine if the child is capable of higher levels of performance.

Unscorable responses, although they do not contribute to the probe score, provide valuable anecdotal information that is helpful in assessing or documenting a below-basal level of performance. In addition, noting the number of different verbs used in a child's responses can be an informal index of a child's verb vocabulary. Use of only one or two different verbs throughout the probe may indicate that the child has an impoverished verb vocabulary. In this case, further evaluation in this area may be warranted.

Just as the sentence context is very important in assessment, if third person singular present tense is selected as a therapy target, it is very important to pay close attention to the sentence context in intervention activities. Simple labeling activities, although very important as a way to increase a child's verb vocabulary, would be an ambiguous context for teaching the obligatory properties of the morphosyntactic rules. Instead you would want to encourage a child to provide complete sentences as responses in activities designed to teach third person singular structures. You may also use the cloze elicitation procedure and provide the subject of the clause, as in the probe procedures. The point is that if a child simply produces the names of activities in a listing of different lexical verbs, this does not constitute a clearly obligatory context for third person singular, and ignores the inherently morphosyntactic nature of the required usage.

Interpreting the Past Tense Probe Results

The Past Tense Probe provides information regarding three elements of past tense usage: past tense for regular verbs, i.e., those that use a /t/, /d/, or /ed/ affix; past tense for irregular verbs, e.g., those that employ stem-internal vowel changes to mark past tense; and overregularizations, in which a child applies the regular *-ed* ending to an irregular verb (e.g., rided, maked). Inclusion of the two types of verbs, regular and irregular, enables the user to learn about a child's understanding of the need to mark verbs for past activities, the likelihood that a child will provide the marker in an adult form, the difference in likelihood attributable to regular rules versus irregular rules for doing so, and the likelihood that a child will mark past tense even if the attempt is in a childlike phonological form.

The preferred elicitation procedure is designed to elicit a spontaneous response from the child (one that includes a past tense verb form within a context of a complete sentence) and to avoid providing a model for the child to imitate. The acceptable alternate elicitation procedure involves the use of a prompt in the form of a cloze procedure. In this case, you provide the subject and the child completes the rest of the response. These procedures ensure that the child's response is based on his or her grammatical system or skills and provides you with a clearer view of the child's capabilities and limitations in this area.

Most tests focus on the semantic elements of "pastness," i.e., a child's awareness of past tense morphemes as an index of a child's general understanding of the concept of "past" in contrast to that of "present." The Rice/Wexler focuses on the obligatory grammatical properties of past tense marking and how that is part of a tense marking system that applies to the present tense (in the third person singular present tense marker and in the present tense forms of *Be* and *Do*). If there is no sign of past tense marking, you should carry out further assessments to determine whether or not a child

has a sense of pastness. This could be apparent, for example, in lexical phrases such as "the other day" and "before now." For children who do not show any evidence of past tense understanding, either morphologically or semantically, an immature concept of "pastness" should be considered as a likely underlying reason for a failure to use past tense on this probe.

In the age ranges evaluated with this instrument, it is likely that the children will use past tense morphology for the items *sometimes*; what changes is that the likelihood of this use increases with age. A child with a language impairment is less likely than his or her age peers to use past tense morphology as a grammatical marker. Given the sometime use of the past tense morphology, the assumption is that children have a sense of pastness (which can be corroborated with other indicators such as semantic phrases), and the problem lies in their understanding of the obligatory properties of the past tense morphology. This assumption is supported by the generally high level of intercorrelations among the probes, as reported in Chapter 4.

A child's performance on the regular verb items is an index of how likely he or she is to mark past tense and follow the regular phonological rules for doing so. A child's performance with irregular verbs captures the obligatory property of tense-marking and captures a child's knowledge of the exceptional phonological properties of marking past tense in those verbs that do not follow the regular phonological rules for past tense morphology. Children who do not grasp the exceptional rules sometimes overregularize, or misapply, the regular morphology rules to irregular verb stems, e.g., they may say "writed" instead of "wrote." Children who overregularize show that they know they need to do something to express past tense; they just do not exactly know how to do so.

The Past Tense Probe offers multiple scores to describe a child's knowledge and use of the past tense. The Past Tense Probe score is an overall past tense score that combines a child's performance on regular and irregular verbs. This score is used to determine whether or not a child is performing at or above criterion in the area of past tense; it is also used in the Screening Test and in calculating the Elicited Grammar Composite.

You may also look at individual performance on regular verbs and irregular verbs. The regular past tense score captures a child's level of obligatory use in the phonologically regular forms. The more complex irregular verbs are described in two scores. First, the Irregular Past Tense score provides information regarding a child's use of irregular past tense verb constructions that follow the adult form (such as "caught" and not "catched"), and second, the Irregular Past Finite Score captures when a child knows that a past tense construction is required, but cannot formulate it correctly. In the Irregular Past Finite score a child is given credit for overregularizing an irregular verb, i.e., "catched" instead of "caught," "writed" instead of "wrote." Growth curves for the three supplemental past tense score calculations (Regular Past Tense, Irregular Past Tense, and Irregular Past Finite) can be found on the *Record Form*.

The discussion here will examine the interpretation of each past tense score. It is important to note that, with only rare exceptions, children's incorrect responses to the attempted past tense items (that is, on-task responses with an appropriate lexical verb) either omit the past tense morphology or misapply the regular past tense rule to irregular forms.

The Regular Past Tense score can be used as an index of a child's progress toward the adult grammar when there are no additional complications of dealing with exceptions to the regular morphology. The Irregular Past Tense score requires further consideration for interpretation. In some interpretations of past tense acquisition, the regular forms are assumed to be generated by a rule-generating procedure. The irregular forms are

assumed to be generated by a system of individually learned lexical items, e.g., "catch" is learned as the present tense form of the verb and "caught" as the past tense form. Although the technical interpretations of this possible generation difference are hotly debated in the scientific literature, the observation that children sometimes generate items such as "*catched" is not controversial and is generally agreed to be a sign of progress toward the adult grammar, not a regression to a more immature grammar. Indeed, the outcomes of the Rice/Wexler indicate that children continue to learn the exceptional properties of irregular verb forms for quite some time.

What is important to capture is the extent to which a child is making progress in mastering the obligatory properties of past tense, regardless of the complexities of the phonological forms of the lexical verbs. This is the reason that two scores are obtained for irregular past tense; the Irregular Past Tense and Irregular Past Finite scores. The Irregular Past Tense score can be interpreted as marking how much progress a child is making on two fronts: knowledge of obligatory properties of tense-marking and knowledge of how irregular lexical verbs show past tense. The Irregular Past Finite score can be interpreted as more directly comparable to the Regular Past Tense score, in that it shows progress toward knowledge of obligatory properties of tense marking.

With regard to clinical applications, each of these scores yields useful guidance for planning intervention activities. Comparison of a child's Irregular Past Tense score and Irregular Past Finite score can help determine when a child understands past tense but is still learning how to form irregular past tense forms. If a child's score on both indices is low, you can assume that the child is not aware of the need to mark past tense consistently. This child probably is generating a high proportion of unmarked (bare-stem) forms of verbs in both the regular and the irregular forms of past tense. This evident pattern would result in a low Past Tense Probe Score when compared to the criterion scores. Such a child should receive intervention focused on the obligatory properties of past tense, with only secondary consideration of the irregular exceptions to the general rules.

Conversely, if a child scores higher on Irregular Past Finite than on Irregular Past Tense, it is likely that he or she is using a number of overregularizations and may be making suitable progress in understanding of the finiteness marking properties of past tense. This could be checked against the child's Regular Past Tense score. If the Regular Past Tense score and the Irregular Past Finite score are comparable and within age expectations, the child may be showing the normal pattern of creative errors on the way to mastery of the irregular past tense. If intervention in this area is pursued, it would be more appropriate to focus on the irregular past tense formation than to focus it on the need to mark past events with a past tense marker. For example, such intervention could consist of practice on individual lexical items that show irregular past tense in a context that alternates present and past tense forms, to help a child to learn the relevant alternations.

As is true for the third person singular morphology, bare-stem responses, such as "paint" for "painted" and "write" for "wrote," are hallmark responses of children with language impairments, particularly, Specific Language Impairment. Bare-stem responses for irregular past tense items also may be helpful in identifying children with language impairments among those who speak Southern White or African American Dialect. Oetting and McDonald (2000) report that bare-stem forms for irregular past tense are a characteristic of children with language impairments among those who speak Southern dialects.

Finally, in reviewing a child's performance on this probe, you should note the number of unscorable responses (including different tenses or sentence structures, or other off-topic responses) the child provided and the number of items to which to the child

did not respond. High numbers of unscorable responses or no responses are possible indicators of immature language that is below the level of this probe, and/or possible behavioral factors that may contribute to diminished performance. Children with a high number of unscorable responses, in combination with only occasional bare-stem forms of lexical verbs, should be evaluated further to determine whether or not they have a concept of past events. Children with a variety of marked past tense verbs combined with a number of unscorable responses or no responses are more likely to have behavioral or situational factors that hamper an estimate of their actual abilities.

Production Probes and A' Values: Implications for General Intervention Strategies

As noted in the previous discussions, the Rice/Wexler probes are designed to provide a detailed analysis of a child's understanding of the grammatical notion of finiteness. Further, performance on one task is thought to be related in a principled way to performance on the other tasks. As reported in the correlational analyses among the probes in Chapter 4, in general the level of performance for a child is likely to be related across tasks, such that as the score for one probe increases relative to that of other children, the score for another probe is likely to do so as well (hence the moderately high positive correlations). This does not mean that all probe scores will be at the same actual level of performance or uniformly mid-high in level. It simply means that relative performance, within the group, is likely to be similar across measures.

The interrelatedness of performance on the probes suggests that there could be value in planning intervention that highlights the entire set of tense markers in a composite package. This way of viewing grammatical training is very different from the more conventional approach of treating individual morphemes separately. Research findings reveal that morphemes are not likely to show equal levels of difficulty for young children with language impairments. The cluster of finiteness markers assessed in the Rice/Wexler are likely to be more difficult than present progressive *-ing* or plural *-s*. The focus on the finiteness morphemes suggests that a new approach to intervention could be to consider the entire set when planning intervention. Although this approach is conceptually reasonable, it has yet to be investigated in formal studies of intervention effects. It may be most useful to explore a strategy in which third person singular *-s*, past tense, and *Be*, and *Do* forms are presented in a mixture of practice items, to determine if a child can grasp the obligatory notion in one or more of the targeted contexts and then spontaneously generalize to one or more of the other contexts. Conversely, if a child shows a pattern of low performance on some (but not all) of the probes, training activities could focus just on the affected morphemes. Perhaps just the third person singular *-s* and the past tense show low performance, in which case it may be useful to teach them together, with the same set of lexical verbs that could sometimes show present tense morphology and sometimes show past tense morphology, e.g., "I jump; I jumped; He jumps; He jumped." With such a set of contrasts a child may be able to deduce the underlying principles. For a child who shows low performance on questions for both *Be* and *Do* forms, it may be useful to teach those forms as an interrelated set, contrasting sentences such as "Does he eat cookies?" and "Is he eating cookies?/Is he hungry?"

Comparisons With Related Test Data

As noted elsewhere in the manual, the best use of the Rice/Wexler is in combination with other assessments. It is essential to know if a particular child has difficulties with multiple dimensions of language, or if the difficulties are more limited, and to evaluate related competencies such as hearing ability and general intellectual functioning. In cases of generalized language impairments, the Rice/Wexler can direct attention to other areas of deficit that may not have been previously detected. It may be, however, that subsequent assessments reveal that performance on the Rice/Wexler indicates the only area of readily identified deficit. It is possible for a child to score within the normal range on an omnibus test and yet have performance below criterion levels on the grammatical markers assessed by the Rice/Wexler. In the case studies encountered in research labs, this scenario appears when a child has relatively high levels of nonverbal intelligence and relatively good performance on semantic dimensions of language (perhaps even with a receptive vocabulary level within normal range). In such cases, a relatively narrow grammatical impairment can nevertheless be a striking academic and social disadvantage. If a child cannot sort out the obligatory properties of tense-marking, he or she is likely to become confused or not perform as well as possible on the numerous reading readiness and early literacy tasks provided in early elementary curricular materials. Unless the source of difficulty is accurately identified, the child could be misunderstood to have a problem with motivation or with attention to tasks. When more detailed analyses are carried out to examine related grammatical structures, such as complex clause formation or complex sentences, more extensive manifestation of grammatical difficulties can appear. These children should be considered candidates for intervention activities in the context of preparing them for school entry and associated reading readiness activities, or, if they are already in school, to better prepare them for the early reading curriculum.

Interpreting Children's Performance Relative to Nonverbal Intelligence or Parent Education Level

Although children's performance levels on language tests are often found to be affected by their performance levels on nonverbal IQ tests, and by parent education level, those two factors are not associated with their performance levels on the Rice/Wexler probes and A' values. As reported in Chapter 5, the Rice/Wexler performance level of the language disordered group (for whom nonverbal intelligence scores are available) correlated at very low levels with nonverbal intelligence, suggesting that low performance on the Rice/Wexler can and does appear in children with levels of nonverbal intelligence within or above normal limits. Conversely, it would not be safe to assume that low levels of performance on the Rice/Wexler is predictive of low levels of performance on nonverbal intelligence assessments. A similiar observation holds for parent education levels, which are not associated with children's performance on the Rice/Wexler test. Children can be below criterion on the grammatical marker regardless of their parents' education levels. The cause of low performance on the grammatical marker is not known, and does not seem to be obviously related to children's general intellectual competencies, within the broad range of borderline-to-normal or above, nor is it related to the general parental resources indexed by parental education levels.

Research and Development 4

History of the Instrument

The initial versions of the tasks included in the *Rice/Wexler Test of Early Grammatical Impairment* were developed as part of a program of research sponsored by an award from the National Institute of Deafness and Communication Disorders to Mabel Rice and Kenneth Wexler, beginning in 1993. The objective of that research was to determine if children with Specific Language Impairment (SLI) differed from children in a control group in their use of finiteness markers, and to track the change in their (the children with SLI) grammar over the period from age 5 years to age 8 years. Because younger children also participated as controls, the lower age range of participants was 3.00–4.11. This program of research is ongoing, and continues to track these children as they approach adolescence. A summary of key outcomes from that program of prior research, and comparisons to the outcomes of the standardization research data for the Rice/Wexler, can be found at the end of this chapter. Suggested reading for other research in this area is also provided.

Probe Development and Previous Testing

The antecedents for the *Rice/Wexler Test of Early Grammatical Impairment* were developed as experimental probes for the program of investigation carried out by Rice and Wexler. Following the literature from previous research studies, the original protocols included spontaneous language samples, the experimental picture elicitation tasks, and the formats for eliciting forms of *Be* and *Do*, and grammaticality judgments. Since the work of Brown (1973), it has been recognized that the percentage correct in obligatory contexts of morpheme use is a useful way to describe young children's acquisition of morphology. Among the advantages are that the technique has strong ecological validity, in that children use their morphology in their own utterances, and the summary measure of percentage correct (for the same child) is quite robust across different situations. At the same time, there are significant limitations as well. The primary limitation is that the method requires a great amount of time for transcription, coding, and data summarization, which often rules it out for clinical application. Another important limitation is that not all morphemes are likely to appear in sufficient numbers to generate a stable calculation of percentage correct. Regular past tense, for example, often appears relatively infrequently in young children's language samples, whereas irregular past tense forms are more likely. This is thought to reflect the preferred verbs of young children (Rice & Bode, 1993; Watkins, Rice, & Moltz, 1993). Other under-sampled forms are questions formed with Be in copula and auxiliary contexts, to parallel their use in declarative contexts; and auxiliary Do questions, even though children's

knowledge of these forms is important theoretically. The development of experimental probes allowed for less time-consuming assessment than spontaneous sampling, observation of grammatical contexts for forms likely to be infrequent in spontaneous samples (regular past tense, *Be* and *Do* in questions), and comparison across spontaneous and elicited probes for the forms that appear in both contexts.

A comparison of the results of spontaneous language samples versus elicited probe tasks showed that that the morphemes under investigation present the same patterns of growth over time in elicited and in spontaneous measurement, and that throughout this age range elicited and spontaneous measures consistently differentiate affected children from the younger control group. Furthermore, statistical analyses with structural equation modeling methods showed that there was no support for the idea that spontaneous measures for a given morpheme "lead" probe outcomes, either within or across times of measurement. This outcome held for the affected children and for the younger children in the control group. Based on these results, development of the experimental probes focused on elicited probe tasks.

Development of the experimental probes (or elicited probe tasks) addressed several issues. Under the direction of the authors, as the items and formats were determined the format for each probe was extensively tested with young children who were developing typically. Item selection was a focal issue for the development of the Phonological, Third Person Singular, and Past Tense Probes. For each of these probes, lexical items were selected because they were likely to be familiar to young children (as indexed by appearance in Hall, Nagy, & Linn's [1984] compilation of spoken words by children ages 4.05–5.00), and could be clearly depicted by line drawings or pictures. Preliminary testing was carried out to ensure that young children could name the pictures with the expected lexical items.

The Phonological Probe consisted of monosyllabic nouns or verbs that ended in the phonemes used in the allophones of third person singular -s (i.e., *s, z*) and regular past tense (i.e., *t, d*). The Third Person Singular Probe items consisted of people pictured at work, and the researcher labeled the pictures according to occupations. The child's task was to describe the activity the person pictured carried out at work, which allowed children to draw upon their own verb vocabulary for selection of appropriate lexical verbs to use as a stem for the targeted *-s* affix. The Past Tense Probe consisted of lexical verb items selected for familiarity, ease of naming, and ease of visual depiction. Items were selected to represent the *-t, -d,* and *-ed* allophonic variants of regular past tense, and internal vowel changes for irregular past tense, with the exception of one irregular item that involved a final consonant change (i.e., *make/made*) and one that involved a vowel change at the beginning of the word (i.e., *eat/ate*). The lexical stems were also selected to have a variety of final consonant types (so no one consonant type, such as an alveolar stop, predominated, in case of final consonant effects on affixation).

Two aspects of assessment are of great importance to the measure of finiteness. One is that there be a sufficient number of items to be able to calculate a percentage of correct responses for a particular form. The way affected children differ from unaffected children is in the probability that they will use a given form. Thus, it is necessary to have enough items to capture the likelihood that they will use a marker. Methods of assessment that are based on a scale of 0 (no use), 1 (one use), or 2 (more than one use), which does not capture the all-important element of probability of use in obligatory contexts, are not as sensitive for detection of grammatical impairment.

The second important element of measurement is the creation of an obligatory context for the use of a given form. Young children are masters at avoiding grammatical contexts that require linguistic specificity for morphological contexts; this means that they can generate responses that are ambiguous because it is not clear what was

intended. The tasks included in the Rice/Wexler are designed to maximize a full clause context where finiteness is unambiguously required in the adult grammar. This is an important element added to the familiar picture elicitation techniques for the Third Person Singular Probe and the Past Tense Probe. Although this task format has been widely adopted, the experimental tasks that appear in the *Rice/Wexler Test of Early Grammatical Impairment* are believed to be the first in formal testing to recognize the need for a full clause obligatory context to evaluate finiteness marking. In this way it is a measure of morphosyntax, not simply morphological affixation.

The Be/Do Probe is an adaptation of an elicited production technique with puppets that has been used in studies of children's question development (Thornton, 1996). The task is designed to elicit questions and declaratives, with singular and plural subjects, for both copula and auxiliary contexts. This design provides a robust number of contexts for the calculation of the probability of marking in each of the contexts. The combination of statement and question contexts for *Be* forms allows for separation of several elements involved in the grammatical system. At the level of declarative sentences, a child must know that a form of Be is obligatory in declarative contexts in both copular and auxiliary uses, with singular and plural subjects (e.g., T*he bug is hungry; the bears are jumping*). At the level of questions, a child must know about the function of asking, and that the form of *Be* is obligatory, and moves to the front of the sentence for singular and plural subjects and copular and auxiliary contexts. A very simple form of question asking is a declarative sentence with a rising intonation (e.g., *The bug is hungry?*). This prosodic manipulation often appears in sentences with deleted forms of Be (e.g., *Bug hungry?*), which is a way for a child to show that asking is understood, but the grammatical rule for obligatory finiteness (i.e., a form of *Be*) and/or the rule for movement of the form of *Be* to the front of the sentence for questions may not be understood. Extensive pilot testing revealed this information is best gathered in a play-like asking/telling descriptive story situation, using a puppet intermediary to establish the necessary referent conditions. The Grammaticality Judgment tasks were developed to provide a direct measure of children's willingness to accept sentences with omitted finiteness markers. Of interest was whether or not children omitted finiteness markers because of some unspecified production limitation, rather than because of an underlying grammatical limitation (Bishop, 1994).

To appropriately evaluate this possibility, the Grammaticality Judgment Probe included items that would help identify whether or not children knew the difference between "Patsy walks" and "Patsy walk," and if they knew that "*Patsy walk" is ungrammatical.

Grammaticality judgment tasks in general are thought to be difficult for young children, and in particular for children with language impairments (Kamhi & Koenig, 1985). One issue is that young children are likely to focus on the semantic elements of sentences and provide judgments based on truth values, instead of on whether or not a sentence is grammatically well-formed (Gordon, 1996; McDaniel & Cairns, 1996). The other issue is that the preferred measurement is an A' measure, which takes into account a child's performance on grammatical and ungrammatical items, and a child's preference for saying "yes" to items. This index requires that for a given A' calculation there be an equal number of grammatical and ungrammatical items. So the Grammaticality Judgment Probe had to focus children's attention on the grammatical, not semantic, elements of a sentence, and it had to be a format that could carry a number of items.

As with the different contexts for the Be/Do Probe, the Grammaticality Judgment Probe items needed to help determine whether or not children would be likely to accept simple sentences with omitted obligatory forms of Be, the past tense, and the third person -s, i.e., whether or not their acceptance of nonfinite clauses would

be evident across the contexts where they were likely to produce nonfinite clauses. Conversely, it was also important to know if their judgments paralleled their correct production, i.e., if they were able to identify grammatical errors that they were unlikely to make in their own productions

The format used in the Grammaticality Judgment Probe is also unprecedented in formal testing. It follows a story description format in which the children observe as the examiner acts out a simple story with small toy objects. Robot toys were introduced as "people from outer space," and the child was asked to tell the examiner if the robots' speech was "good" or "not so good."*

This format was interesting to young children, and eliminated spurious truth-value judgments. This seems to be because the story description puts the event frame, the reference frame, and the speech frame in alignment (i.e., the statements are about immediate events that are observed and understood by the child). Practice items proved to be successful in orienting the children to the grammatical focus of the judgments.

** During the pilot testing, little girls were reluctant to tell the robots their speech was "bad," so "not so good" was used instead.*

Tryout Research

The Tryout version of the *Rice/Wexler Test of Early Grammatical Impairment* included some changes and additions to the previous research versions. Several stimuli from the original version of the Phonological Probe were substituted and a total of seven new stimuli were added. These changes were made to determine the group of stimuli for each phoneme that elicited the pictured stimuli (target word) most consistently for all ages of children and for children from all geographic regions.

Five additional items were added to the Tryout version of the Third Person Singular Probe. These additional items were included to ensure that the items appealed to the full age span of the Tryout instrument; to ensure that the occupations depicted were familiar to children of all race, geographic, and socioeconomic backgrounds; and to obtain data on a larger number of items than would ultimately be needed. This would enable the number of items in the standardization version to be reduced to the fewest number of items needed to maintain the integrity of the instrument and to definitively differentiate between individuals who have a language disorder and individuals who have normal language skills.

Nine items were added to the Tryout version of the Past Tense Probe. Four regular past tense items were added in order to have an equal number of each regular verb ending (-d, -t, -Id): three final /-Id/ items and one final /-d/ item. Five irregular past tense items also were added to obtain data on a larger set of irregular items than would ultimately be needed for the final version. This enabled the test developers to select the items that appeared to work the best in terms of appealing to a wide group of children and in terms of discriminating between normal and disordered performance.

Tryout research of the *Rice/Wexler Test of Early Grammatical Impairment* was conducted during the spring and summer of 1999 by 75 professionals, including speech-language pathologists, early childhood educators, educational diagnosticians, and psychologists. Each examiner completed a background questionnaire and a practice case before being approved to test additional cases. Trained staff at The Psychological Corporation reviewed each practice case to ensure that the case was administered correctly and that the responses were being recorded accurately. Throughout the Tryout research phase, examiners were provided feedback regarding administration and scoring procedures,

and additional instruction on recording responses. This feedback was provided via telephone, newsletters, and e-mail.

The data from the Tryout research were analyzed to determine which items performed best in terms of eliciting the grammatical structure being targeted. Items that were identified as eliciting the targeted structure with less frequency were examined to identify possible causes. The following possible causes were identified: the artwork was problematic or distracting to the children; an occupation or activity represented was unfamiliar to some children; and a vocabulary term used in the stimulus was unfamiliar or vague. In addition, some items were found to be less effective because they represented potential gender, regional, or cultural bias. When possible, items found to be problematic were discarded before the standardization phase began.

The Tryout testing was completed on a total of 367 children including: 146 children between the ages of 3.00 and 6.11 whose language skills were considered to be normally developing or who had no known language impairments; 101 children between the ages of 4.00 and 8.11 whose language skills indicated that the child met the criteria for Specific Language Impairment; and 120 children between the ages of 3.00 and 8.11 whose language skills indicated that the child met the criteria for Non-Specific Language Impairment.

Table 4.1. shows the Tryout sample by age and language status.

Table 4.1 Tryout Research Sample by Age and Language Status

Age	Normal Language Group	Specific Language Impairment	Non-Specific Language Impairment
3.00–3.05	12	0	2
3.06–3.11	19	0	5
4.00–4.05	17	8	5
4.06–4.11	27	8	13
5.00–5.05	21	11	3
5.06–5.11	15	14	16
6.00–6.11	35	22	19
7.00–7.11	not tested	23	26
8.00–8.11	not tested	15	31
Total	146	101	120

Bias Review

The presence of bias in standardized tests is undesirable, not only because it does not take into account individual differences, but because it can result in inaccurate scores or results. For example, items containing regional expressions that only some students use, or items that require background knowledge or information that only some students have, can lead to an unfair assessment of a student's actual skills. Likewise, a language test that evaluates specific morphological structures that may be considered optional for some test takers would be considered biased against those students. To the extent possible, and when appropriate, test bias is eliminated during the development process.

The Rice/Wexler is unique in its design and purpose. As specified in the purpose statement, this instrument is designed for, and its validity and reliability can only be assured when it is used with, children who speak Standard American English. For all other populations, use of this instrument must be considered with caution. The interpretation of results may be greatly influenced by a child's dialect or native language if it is not Standard American English.

To ensure that test bias was eliminated or reduced for the intended populations, the Tryout edition was submitted to a panel of speech-language pathologists who have expertise in multicultural and/or minority issues. The panel reviewed test items, administration directions, stimulus pictures, and the manipulatives for potential gender, race/ethnicity, class, cultural, and regional bias.

The reviewers confirmed much of what was already known about the instrument. Potential bias related to gender, class, or region presented few significant concerns about the instrument. Overall, the picture stimuli and the vocabulary used in the items were found to be generally appropriate for most children in the United States who speak Standard American English. Reviewers frequently commented that a particular word used in an item may not be familiar to children in their area or from a particular socioeconomic status. In most of these cases, the reviewer commented that having the associated picture helped, as did the fact that the examiner can model the target word on the Phonological Probe and name the occupation or activity being represented in the Third Person Singular Probe and the Past Tense Probe. When possible however, these items were revised or deleted from the instrument.

A review of the instrument was also solicited from potential users in the United Kingdom, Australia, and Canada to evaluate the possible interest or usefulness of the product in these areas. In cases of international review, more concerns regarding bias were evident. Particularly, specific vocabulary used may not be familiar, and picture representation of sports and other leisure activities may not be as familiar, although it appears that the structures being tested would be appropriate and of interest in assessment. Because the target population for the instrument is children who speak Standard American English, and because it would have required a substantial revision of the instrument, eliminating potential international bias was secondary to addressing concerns for children within the United States. However, users in these areas who are interested in using the instrument are invited and encouraged to do so, with the caution that they should ensure it meets the needs of assessment for their populations.

As expected, reviewers who provided information regarding children who do not speak Standard American English (including African-American English, 2nd language influenced English, geographical/regional dialects such as Southern White English, Appalachian English) reported that these children are likely to experience difficulties with much of the content included in this instrument as a result of the dialect or geographical influences, as opposed to necessarily resulting from an actual language disorder.

Summary of the information obtained from bias reviews:

Potential bias relative to children who speak African American English (AAE)—Optional deletion of third person singular -s is a primary characteristic of AAE subject-verb agreement (e.g., "*he drive the car"). Regular past tense forms are optionally included (e.g., "*yesterday he watch T.V. before school") and the forms targeted in the Be/Do Probe may be optionally included or may not agree in number with the verb (e.g., "*This my backpack;" "*They is happy"). In addition, throughout the sections of the instrument there are items tested that could be scored with an AAE speaker as *zero -ing* (e.g., "*the bear is sit"), and *zero plural* (e.g., "*these are two spoon"). Because of the option for inclusion, exclusion, or agreement of these morphemes for AAE speakers, the effect of these dialectal differences on the outcomes or results of the instrument is not fully known.

Potential bias relative to children who speak or whose speech is influenced by a Spanish language—Optional deletion of -s and subject-verb agreement errors may be characteristic of Spanish language influence (e.g., "*doctors examines" or "*doctor examine"). For the Past Tense Probe, omission of -ed for regular verbs or use of present tense (e.g., "*Yesterday he cry") may indicate influence of Spanish language. This influence may also be evident in the Be/Do and Grammaticality Judgment Probes, where subject-verb agreement errors are likely, or acceptance of statements or questions with subject-verb agreement errors are likely.

Potential bias relative to children who speak or whose speech is characteristic of Asian-influenced English—The phonemes targeted in the Phonological Probe and subsequently used as morphological endings throughout the instrument may not appear or may be omitted in some Asian languages. The Past Tense Probe responses would also be affected because lexical verbs do not change for tense in some of these languages. Likewise, the concepts of Be and Do do not exist in some Asian languages.

For these reasons, the *Rice/Wexler Test of Early Grammatical Impairment* would be inappropriate for diagnosing a language disorder in children unlike the children included in the standardization research. More research is needed to clarify possible influences or test bias due to cultures, dialects, or second language influence on English to fully understand the usefulness of the Rice/Wexler with children included in these populations. It is extremely important that you use this instrument for diagnostic purposes only with children for whom the test was developed, i.e., children who speak Standard American English. On the other hand, the results of the test can provide descriptive information about a particular child's progress toward the standard English grammar system of finiteness marking, which could be helpful in clinical and educational settings.

Table 4.2 lists the bias panel participants for the Rice/Wexler.

Table 4.2 Rice/Wexler Bias Panel Members

Li-Rong Lilly Cheng, Ph.D. Assistant Dean Global Program Development San Diego State University	Henriette W. Langdon, Ed.D., CCC-SLP Associate Professor Special Education San Jose State University
Christine Vining, M.S., CCC-SLP Associate Director Health Sciences Center The University of New Mexico	Julie Washington, Ph.D. Senior Associate Research Scientist Institute For Human Adjustment University of Michigan
Toya A. Wyatt, Ph.D. Associate Professor Dept. of Communication California State University, Fullerton	

Examiner Input and Feedback

All examiners who participated in the Tryout field-testing were asked to complete a questionnaire that evaluated the clarity and usefulness of the test directions, test items, visual stimuli, manipulatives, and the *Training Videotape*. The feedback from examiners indicated a need for further clarification regarding using the prompts and recording responses. Most questions and comments regarding using the prompts related to the Third Person Singular Probe and the Past Tense Probe, especially questions dealing with supplying prompts or supplying a subject for the child. Several comments indicated children's confusion with the artwork on a few items and some regarded children's reactions to the manipulatives. Examiners also expressed concerns that the instrument was too long. Based on this feedback, some changes and refinements were made to the instrument for the Standardization edition.

Analyses

All Tryout research *Record Forms* were reviewed by trained staff, and responses to test items were scored by trained personnel. Statistical analyses were conducted to evaluate how well the test discriminates between children with and children without language disorders, and to determine whether or not the test could be shortened without compromising the integrity of the instrument. Based on the combined results of statistical analyses, recommendations and comments obtained from bias panel reviewers, and examiner feedback, some items were deleted and revisions and refinements were made to the administration procedures.

Standardization Research

The Standardization and related validity and reliability research for the R*ice/Wexler Test of Early Grammatical Impairment* took place during the fall of 2000. The instrument was standardized with 393 children between the ages of 3.00 and 6.11, whose language skills were considered to be developing normally, and 444 children between the ages of 3.00 and 8.11 who had a diagnosed language disorder.

To participate in the Standardization research each child had to

- be from a home where English is spoken at least 75% of the time,
- speak Standard American English (SAE),*
- have adequate hearing and vision,
- be able to take the test in English in a standardized fashion, and
- pass the Phonological Probe.

** By Standard American English (SAE), we mean that the child speaks English that is considered to be a widely, socially accepted variety of English that is relatively unmarked with respect to regional or, in some cases, dialectal characteristics of English. Children whose English is influenced by a second language that is spoken either in the home or by the child, would not be considered to speak SAE. Also, children who speak a dialect such as African American English or Appalachian English, where some standard English grammatical structures may not be required, would not be considered to speak SAE.*

Selection and Qualification of Examiners

One hundred and eighty professionals participated in the Standardization research and the validity and reliability research. They included speech-language pathologists, early childhood educators, educational diagnosticians and psychologists. Each examiner completed a background questionnaire before being approved to participate in the study. Trained staff at The Psychological Corporation reviewed the first case each examiner completed to ensure that the child tested was appropriate for the study and that the examiner had administered the test and recorded the responses correctly. Each examiner was provided specific feedback regarding any areas of the test case that were problematic. Each additional test was also reviewed to ensure the accuracy and validity of the data. Throughout the Standardization phase, examiners were provided feedback regarding administration procedures or scoring or recording responses. This feedback was provided via telephone, newsletters, and e-mail.

Description of the Standardization Sample

The *Rice/Wexler Test of Early Grammatical Impairment* standardization research was conducted with two groups of children; a group of children without known language disorders (normal language group) and a group of children with known language disorders (language disorder group). The sample of children in the normal language group was stratified on the basis of age, gender, race/ethnicity, geographic region, and parent education level. The sample of children in the language disorder group was stratified on the basis of age only.

Description of Children in the Normal Language Group

The requirement for inclusion in the normal language group was that, in addition to meeting the criteria for all children in the Rice/Wexler standardization research study, the child had not been diagnosed with a language disorder at the time he or she was selected and tested for this study. Clinicians were asked to select children for participation that did not have a known language disorder and who were not suspected of having a language disorder. Following conventional test development procedures, no testing was conducted or required to ensure that these children had "normal" language skills. Inclusion in this study was based primarily on a clinician's judgment that a child was appropriate for the study. It is possible that some children in this group may have language deficits that have either not been investigated or have not been detected through previous testing. Epidemiological assessment of a large sample of kindergarten children found that 7% of the children met a psychometric definition of SLI, and of those children, 29% of the parents reported they had been notified that the child had a speech or language problem (Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997). It cannot be ruled out that among the children classified as "normal" there are false positives, i.e., children not detected as affected. The effect of this possible source of bias is to include lower levels of performance in the "normal" group than may exist in a sample of children whose test scores place them within or above normal range. This information should be taken into account when interpreting the scores obtained relative to the criterion scores developed for this instrument.

Children were not excluded from the normal language group if they were receiving special education services or services as gifted and talented. As a result, 2% of the children included in the normal language group consisted of children who were classified as having attention deficit disorder (ADD), developmental delay, learning disability, speech delay or disorder, or other health impairments. Because one of the qualifications for inclusion in the sample was the ability to attend to and take the test in the standard fashion without modifications, some children with severe disabilities were excluded.

Description of Children in the Language Disorder Group

The requirement for inclusion in the language disorder group was that, in addition to meeting the criteria for all children in the standardization research study, the child had been diagnosed with a language disorder at the time he or she was selected and tested for this study. Clinicians were asked to select children for participation from their caseloads or who were known to the clinician as having a language disorder. The clinicians were required to provide documentation of the language testing that was used to diagnose the language disorder. It was requested that the qualifying test scores be no older than 12 months; however children were accepted into the study with language scores as old as 15 months. The children could also have other conditions in addition to the language disorder, such as a learning disability or AD/HD or an articulation, voice, or fluency disorder. In accordance with the requirements of this study, some children may have been included in this study who had been in language therapy or treatment and might have made significant progress such that if testing was completed today, the child may no longer qualify for the study. Other children may have been included in the study as a result of low performance on omnibus tests for reasons of low vocabulary or deficits in other areas of language that may not result in low performance on the grammatical markers tested on the Rice/Wexler. It is important to note this and consider this when interpreting the Rice/Wexler results.

Also consider that when one looks at the relatively broad, and clinically realistic, criteria utilized in forming the normal and language disorder groups, there is an unknown element of possibly misclassified children. This element is likely to work in the direction of weakening the sensitivity and specificity outcomes. As reported, the outcomes are robust, even with this possible element working against clear differentiation of affected and non-affected children. The evidence presented here is intended to help you both to compare a given child's performance to each group and to keep in mind the composition of the groups when arriving at a decision of affectedness.

Tables 4.3 and 4.4 show the distribution of the standardization sample by age for children in the normal language group and for children in the language disorder group, respectively.

Table 4.3 Rice/Wexler Standardization Sample by Age—Children in the Normal Language Group

Age (years.months)	**n**
3.00–3.05	43
3.06–3.11	50
4.00–4.05	50
4.06–4.11	50
5.00–5.05	50
5.06–5.11	50
6.00–6.05	50
6.06–6.11	50
Total	393

Table 4.4 Rice/Wexler Standardization Sample by Age—Children in the Language Disorder Group

Age (years.months)	**n**
3.00–3.05	20
3.06–3.11	24
4.00–4.05	50
4.06–4.11	50
5.00–5.05	50
5.06–5.11	50
6.00–6.05	50
6.06–6.11	50
7.00–7.11	50
8.00–8.11	50
Total	444

Tables 4.5 and 4.6 show the distribution of the standardization sample by gender for children in the normal language group and for children in the language disorder group, respectively.

Table 4.5 Rice/Wexler Standardization Sample by Gender—Children in the Normal Language Group

	Ages 3.00–5.11		**Ages 6.00–6.11**	
Gender	**n**	**Sample %**	**n**	**Sample %**
Female	149	51	51	51
Male	144	49	49	49
Total	293	100	100	100

Table 4.6 Rice/Wexler Standardization Sample by Gender—Children in the Language Disorder Group

	Ages 3.00–5.11		**Ages 6.00–8.11**	
Gender	**n**	**Sample %**	**n**	**Sample %**
Female	80	33	91	46
Male	164	67	109	54
Total	244	100	200	100

Tables 4.7 and 4.8 show the distribution of the standardization sample by race/ethnic distribution for children in the normal language group and for the children in the language disorder group, respectively. Each child in the sample was categorized by his or her parents as belonging to one of the race/ethnic groups listed.

Table 4.7 Rice/Wexler Standardization Sample by Race/Ethnicity—Children in the Normal Language Group

	Ages 3.00–5.11			Ages 6.00–6.11		
Race/Ethnicity	**n**	**Sample %**	**U.S. Population %**	**n**	**Sample %**	**U.S. Population %**
African American	38	13.0	16.0	4	4.0	16.4
Hispanic	47	16.0	17.4	11	11.0	15.9
Other	13	4.4	5.3	13	13.0	5.1
White	195	66.6	61.3	72	72.0	62.7
Total	293	100	100	100	100	100

Table 4.8 Rice/Wexler Standardization Sample by Race/Ethnicity—Children in the Language Disorder Group

	Ages 3.00–5.11		Ages 6.00–8.11	
Race/Ethnicity	**n**	**Sample %**	**n**	**Sample %**
African American	25	10.3	36	18.0
Hispanic	20	8.2	17	8.5
Other	12	4.9	22	11.0
White	187	76.6	125	62.5
Total	244	100	200	100

Tables 4.9 and 4.10 show the distribution of the standardization sample by geographic region of the United States. The regions are defined by the U.S. Bureau of the Census (1999). Current population survey, March 1999.

Table 4.9 Rice/Wexler Standardization Sample by Region—Children in the Normal Language Group

	Ages 3.00–5.11			Ages 6.00–6.11		
Region	**n**	**Sample %**	**U.S. Population %**	**n**	**Sample %**	**U.S. Population %**
Northeast	60	20.5	18.1	13	13.0	18.7
North Central	67	22.9	23.6	31	31.0	23.9
South	76	25.9	33.0	26	26.0	33.1
West	90	30.7	25.3	30	30.0	24.2
Total	293	100	100	100	100	100

Table 4.10 Rice/Wexler Standardization Sample by Region—Children in the Language Disorder Group

	Ages 3.00–5.11		Ages 6.00–8.11	
Region	**n**	**Sample %**	**n**	**Sample %**
Northeast	32	13.1	32	16.0
North Central	86	35.3	42	21.0
South	62	25.4	73	36.5
West	64	26.2	53	26.5
Total	244	100	200	100

Tables 4.11 and 4.12 report the standardization sample by parent education level. Parent education level was obtained by asking parents/guardians to specify the highest grade completed. If a child's parents had different education levels, the primary caregiver's education level was used. The primary caregiver is considered to be the parent who spends the most time with the child.

Table 4.11 Rice/Wexler Standardization Sample by Parent Education Level—Children in the Normal Language Group

	Ages 3.00–5.11			Ages 6.00–6.11		
Years of Education	**n**	**Sample %**	**U.S. Population %**	**n**	**Sample %**	**U.S. Population %**
11 or less	39	13.3	17.6	7	7.0	17.4
12	49	16.7	31.2	22	22.0	32.2
13–15	87	29.7	28.6	34	34.0	28.9
16 or more	118	40.3	22.7	37	37.0	21.4
Total	293	100	100	100	100	100

Table 4.12 Rice/Wexler Standardization Sample by Parent Education Level—Children in the Language Disorder Group

	Ages 3.00–5.11		Ages 6.00–8.11	
Years of Education	**n**	**Sample %**	**n**	**Sample %**
11 or less	28	11.5	36	18.0
12	77	31.6	72	36.0
13–15	78	32.0	61	30.5
16 or more	61	25.0	31	15.5
Total	244	100	200	100

Scores

The scores obtained from the Rice/Wexler represent a "percent correct" of attempted items within a given item set. Although children with language impairments sometimes use the targeted tense markers, they differ from normal children in the likelihood that they will use the grammatical markers in obligatory contexts. Thus the probes are designed to measure percentage of marker use when required.

Table 4.13 shows the means and standard deviations for each probe score and for the Elicited Grammar Composite for children ages 3.00–6.11 in the standardization research study for the children in the normal language group. Table 4.14 shows the means and standard deviations for each probe score and for the Elicited Grammar Composite for children ages 3.00–8.11 in the standardization research study for the children in the language disorder group. Figures 4.1–4.8 show graphical representation of these data in the form of box and whiskers plots.

For these data, starting with the youngest group of children in the normal language group, the values of the means for the Third Person Singular, Past Tense, Be, Do, and Elicited Grammar Composite are in the range of .60 to .72. With age, these values rise and ultimately settle in the .90–.97 range at the oldest age groups. Also, the standard deviations (the index of within-group variance across individuals) steadily shrink as the age of the children increases. This pattern indicates that, as expected, children approach two things simultaneously: the expected high levels of probability of use of the target morphemes, and uniformly high levels for the unaffected children. This progression is clearly shown in the figures, where the width of the box shows the range of scores between the 25th and 75th percentile. With increasing age, the width of the boxes for the normal group becomes smaller, and the means (roughly, the center of the box) increase and then seem to plateau around 95%.

In comparison, the means for these probe scores for the children with language disorders also clustered in the same range of values, although this range is considerably lower than that of the normal group. At the youngest age level, the means range from .14 to .36, roughly 30–55 points below the normal group, a difference of more than one standard deviation lower than the normal performance levels of the normal group. This suggests that within each age level there is little overlap of the performance level of individuals within the normal group and individuals within the language disorder group. This can be seen by the fact that in the box and whiskers plots there is little to no overlap of the boxes for the two groups. The non-overlapping nature of the distributions of individuals within the groups can also be seen in the calculations of sensitivity and specificity reported (see Appendix B). Finally, although the standard deviations (width of the boxes) decrease with age for the normal language groups, the language disorder groups show a fairly steady standard deviation of roughly .20 to .35 over all the age groups. This pattern suggests that the variation between individuals within the language disorder group is fairly steady during this developmental period. Stated another way, within a given age level of the language disorder group, there is likely to be somewhat more diversity between individuals than for a given normal group, and this is even more striking as the children get older.

Table 4.13 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Normal Language Group

		Third Person Singular Probe		Past Tense Probe		Be/Do Probe (Be Score)		Be/Do Probe (Do Score)		Elicited Grammar Composite		Dropped Marker*		Agreement*		Dropped -ing*	
Age	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3.00–3.05	43	.71	.30	.65	.25	.72	.28	.60	.35	.67	.22						
3.06–3.11	50	.80	.27	.78	.22	.86	.17	.71	.36	.79	.17						
4.00–4.05	50	.87	.24	.84	.19	.87	.18	.74	.34	.83	.20	.70	.25	.75	.25	.73	.31
4.06–4.11	50	.91	.18	.90	.10	.90	.16	.83	.23	.89	.11	.75	.21	.81	.20	.83	.22
5.00–5.05	50	.93	.11	.88	.12	.93	.12	.87	.18	.90	.10	.80	.20	.84	.21	.85	.24
5.06–5.11	50	.97	.06	.93	.08	.93	.08	.83	.20	.92	.08	.83	.18	.87	.16	.87	.17
6.00–6.05	50	.97	.07	.93	.08	.96	.06	.90	.13	.94	.06	.92	.12	.94	.12	.94	.15
6.06–6.11	50	.96	.07	.94	.06	.96	.06	.90	.14	.94	.06	.93	.09	.98	.05	.97	.07

* As described in text, these scores represent A' calculations. All other scores represent a percentage correct.

Table 4.14 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Language Disorder Group

		Third Person Singular Probe		Past Tense Probe		Be/Do Probe (Be Score)		Be/Do Probe (Do Score)		Elicited Grammar Composite		Dropped Marker*		Agreement*		Dropped -ing*	
Age	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3.00–3.05	20	.29	.34	.36	.31	.23	.37	.14	.33	.25	.20						
3.06–3.11	24	.26	.35	.30	.25	.40	.35	.09	.24	.26	.21						
4.00–4.05	50	.38	.34	.38	.30	.48	.32	.21	.32	.36	.24	.43	.30	.46	.35	.45	.38
4.06–4.11	50	.39	.34	.48	.27	.57	.34	.20	.31	.41	.23	.43	.29	.50	.34	.53	.38
5.00–5.05	50	.47	.36	.44	.31	.46	.31	.25	.31	.41	.26	.53	.22	.60	.25	.64	.30
5.06–5.11	50	.47	.35	.49	.30	.60	.25	.30	.28	.47	.24	.58	.20	.65	.25	.69	.29
6.00–6.05	50	.57	.37	.60	.24	.59	.28	.36	.35	.53	.24	.58	.26	.65	.29	.61	.34
6.06–6.11	50	.57	.34	.58	.28	.62	.28	.44	.32	.55	.25	.63	.23	.73	.20	.71	.27
7.00–7.11	50	.69	.34	.76	.23	.79	.23	.67	.31	.73	.21	.76	.20	.82	.20	.84	.24
8.00–8.11	50	.73	.35	.78	.24	.78	.22	.67	.27	.74	.22	.83	.16	.88	.15	.92	.14

* As described in text, these scores represent A' calculations. All other scores represent a percentage correct.

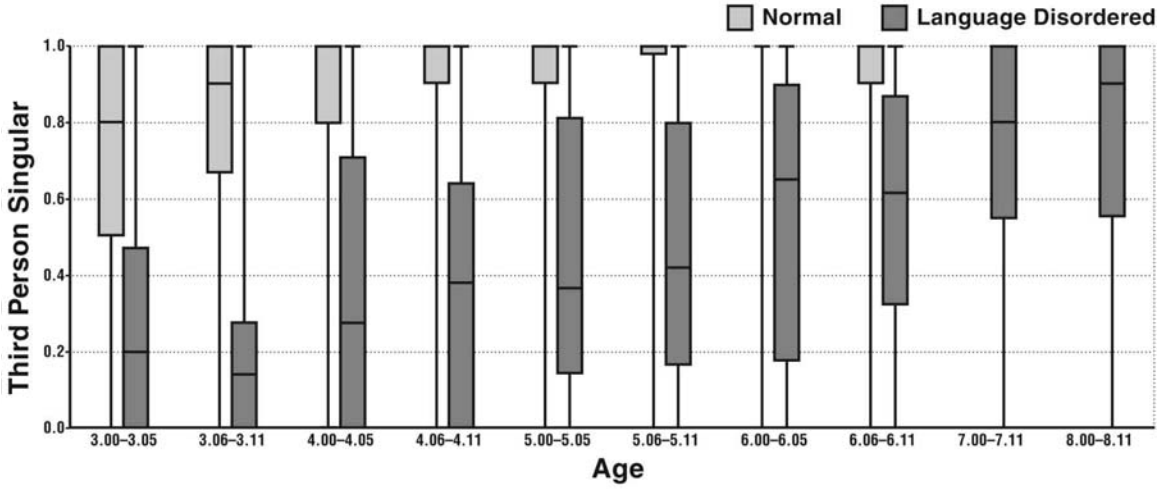


Figure 4.1 Box and Whiskers Plot for Third Person Singular

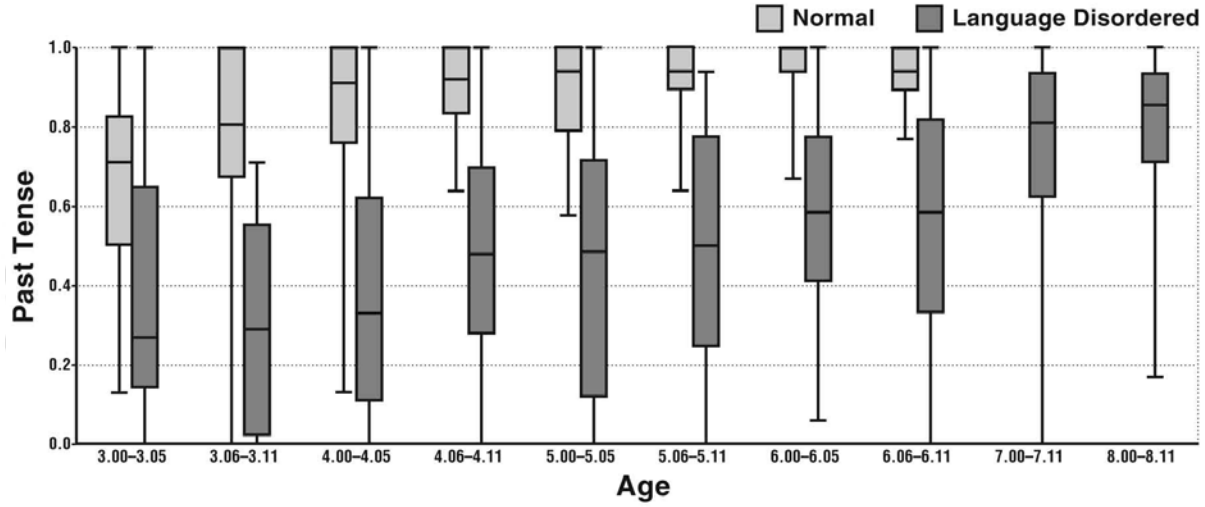


Figure 4.2 Box and Whiskers Plot for Past Tense

Development of Criterion Scores

Criterion scores are provided for the Elicited Grammar Composite and for individual probe scores, with the exception of the Phonological Probe, which is scored as pass/fail only. The criterion scores are used to determine if the child is functioning like those children who have a language disorder or if he or she appears to be functioning like those children who do not have a language disorder. The criterion scores were developed by the authors using their expertise and experience and based, in part, on sensitivity and specificity data collected.

Additional information regarding sensitivity and specificity are also provided in this manual. For each probe score, sensitivity and specificity levels are provided for all possible scores for the probe (0–100). These data are located in Appendix B, where you can examine the range of scores associated with a given level of sensitivity. For example, a score with a sensitivity of .80 means that if a child has a language impairment, there is an 80% chance that he or she will be so identified. This is a way to place a particular child's performance relative to those children of the same age who were known to have language impairments. It would also be possible to place that same score in the normal group of children, to know if the specificity is equally as high. For example, a given score could be where 80% of the children with language impairments scored at or below, and also where 80% of the normal group of children scored above. This score would then sit at the intersection of the two distributions of children, as shown in Figure 1.2 in Chapter 1 for the bimodal distribution.

Growth Curves

For the distribution of probe scores, the mean, the 25th percentile, and the 75th percentile for the children in the normal language study group were calculated for each probe and for the Elicited Grammar Composite for each age group. These values are presented in Table 4.15 and are also provided in graphical format in the form of growth curves. The growth curves for each probe score and for the Elicited Grammar Composite appear on the *Record Form*.

Using growth curves enables you to compare the results of a child's performance to that of his or her age peers (in the normal language group). Another way to interpret the growth curves for the Rice/Wexler is to compare a child's performance relative to his or her progress towards the adult grammar.

Table 4.15 25th Percentile, Mean Score, and 75th Percentile for each Rice/Wexler Probe and for the Elicited Grammar Composite—Children in the Normal Language Group

	Third Person Singular Probe			Past Tense Probe			Be/Do Probe (Be Score)			Be/Do Probe (Do Score)			Elicited Grammar Composite			Dropped Marker			Agreement			Dropped -ing		
Age	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile	25th Percentile	Mean	75th Percentile
3.00–3.05	.50	.71	1.00	.50	.65	.83	.61	.72	.96	.33	.60	.88	.50	.67	.83									
3.06–3.11	.66	.80	1.00	.67	.87	1.00	.81	.86	1.00	.55	.71	1.00	.64	.79	.94									
4.00–4.05	.80	.87	1.00	.76	.84	1.00	.79	.87	1.00	.55	.74	1.00	.73	.83	.99	.50	.70	.90	.56	.75	1.00	.50	.73	1.00
4.06–4.11	.89	.91	1.00	.83	.90	1.00	.86	.90	1.00	.72	.83	1.00	.84	.89	.96	.56	.75	1.00	.71	.81	1.00	.66	.83	1.00
5.00–5.05	.89	.93	1.00	.79	.88	1.00	.90	.93	1.00	.75	.87	1.00	.82	.90	.99	.72	.80	1.00	.79	.84	1.00	.79	.85	1.00
5.06–5.11	.98	.97	1.00	.89	.93	1.00	.87	.93	1.00	.73	.83	1.00	.86	.92	.98	.74	.83	1.00	.80	.87	1.00	.79	.87	1.00
6.00–6.05	1.00	.97	1.00	.94	.93	1.00	.95	.96	1.00	.82	.90	1.00	.91	.94	.99	.89	.92	1.00	.93	.94	1.00	1.00	.94	1.00
6.06–6.11	.90	.96	1.00	.98	.94	1.00	.94	.96	1.00	.90	.90	1.00	.92	.94	.99	.89	.93	1.00	.99	.98	1.00	1.00	.97	1.00

Previous Studies and Comparisons to Rice/Wexler Outcomes

The program of research carried out by Rice and Wexler, and their colleagues, investigated the morphosyntactic development of children with Specific Language Impairment and of children in control groups. Although children with SLI are not the only children with language impairments, they are of special interest because there is no obvious cause for their language impairment and they often go undetected. The search for a grammatical marker was driven, in part, by the assumption that a grammatical marker that would identify these children would also be likely to identify other children with language impairments. Also, the work has investigated whether or not the grammatical marker is an area of language where the affected children perform even lower than younger children at the same general level of language development; that is, whether the grammatical marker is not only slow to emerge relative to age expectations but is even slower to mature than the other areas of language. In general, the findings show that the grammatical marker is slower to mature than other areas of language, which are delayed relative to age expectations. Because of the selective properties of immaturity of the grammatical marker, it is more likely to lag behind and be detected as part of a language impairment in affected children, even children with SLI.

Before summarizing the key outcomes, it is important to describe the clinical criteria for the samples of SLI children who participated in the studies. The children were drawn from the caseloads of speech-language pathologists and entered a longitudinal study in the year before kindergarten, when they were around 5 years of age. Their nonverbal IQ was within normal range; their hearing was within normal limits or above; they were not diagnosed as having sociobehavioral deficits; their speech performance met the criteria for the Phonological Probe; and, finally, they had no evidence of neurological conditions that would be likely to impact speech and language. Their language limitations were the following: they achieved a score of one standard deviation or more below the mean on an omnibus language test and on a separate standardized test of receptive vocabulary; their mean length of utterance was one standard deviation or more below the group means reported by Leadholm & Miller (1992). Thus, the affected children in the longitudinal study would be described as receptive/expressive language impaired without other developmental deficits at the outset of the study. In subsequent studies, described below, affected children were evaluated who showed different diagnostic criteria for SLI and non-SLI language impairments.

The experimental design of many of the studies reported here involves a comparison of three groups: an affected group, a control group of children of the same chronological age, and a control group of younger children with an equivalent length of utterance, generally referred to as a "language matched" group. The interpretive significance of the younger control group is that if the affected children perform at a level lower than their age comparisons on a grammatical marker, this indicates that they are not at age expectations, which could perhaps be true of other elements of their language acquisition, as well. The younger controls enable us to examine whether or not the affected group performs below children at the same general level of language development, i.e., at the same utterance length. If there are differences between the affected group and the language-matched control group, these differences would suggest that the grammatical marker taps into an area of deficit that exceeds that which would be expected of a generally immature language.

The outcomes are summarized here in terms of the previously published findings from the experimental research, and compared to the current Rice/Wexler test results. The citations for the original research reports are provided for the reader who wishes to examine the details.

1. At the age of school entry, children with SLI perform below control groups for every morpheme in the experimental probe set, both on spontaneous measures and on experimental probes (Rice, Wexler, & Cleave, 1995 and Rice & Wexler, 1996).

   As noted previously in this manual, the finding that the elicitation probes detect differences between groups as consistently as the spontaneous measures meant that the test development could focus on the less time-consuming elicitation probes. The test outcomes provide clear evidence of high levels of sensitivity and specificity for the grammatical marker. This is very important new information because the control groups are more inclusive in the test standardization samples than in the earlier research studies (e.g., including children with AD/HD and speech impairments) and the language disorder groups are more inclusive in the test standardization (e.g., including children with language impairments that have expressive deficits only and including children across a wide range of nonverbal IQ levels).

2. Growth in grammatical tense-marking is much slower for affected children than for either of the control groups, although the growth trajectory is similar (Rice, Wexler, & Hershberger, 1998).

   The experimental research outcomes are for the same group of children, studied over time. The test outcomes are for different samples of children for each age group. A series of box and whiskers plots clearly illustrates that for each of the test probes the language disorder group of children's performance falls below that of their age peers. The figures also show that the variation within the normal groups becomes smaller as the children get older, and there is less variation in general within the normal groups than within the language disorder groups (see Figures 4.1–4.8).

   A series of two-way univariate ANOVAs were carried out for each of the test probes, with group and age as between-subjects factors. The outcomes were the same for each probe and the Elicited Grammar Composite: there are significant group effects for each, with the language disorder groups performing below the normal groups; there are significant age effects for each, with the older groups performing better than the younger groups; and there are no significant interactions of group by age on any of the measures. Table 4.16 reports the outcomes of this analysis. The lack of significant interactions indicates that the language disorder groups shadow the changes in the normal groups throughout the time period, i.e., as the normal group improves performance, so does the language disorder group, but the language disorder group does not close the gap. A further follow-up series of t tests were carried out to confirm that the disorder group's performances are below that of the normative group for each measure, at each age level. See Table 4.17 for the outcomes of the t test.

Table 4.16 Analysis of Variance for Rice/Wexler Probes and for the Elicited Grammar Composite

Third Person Singular

Source	df	SS	MS	F	Significance
Group	1	31.71	31.71	447.23	<.0001
Age	5	1.71	0.34	4.81	0.0003
Group by Age	5	0.34	0.07	0.97	0.4300
Error	588	41.69	0.07		
Total	599	75.46			

Past Tense

Source	df	SS	MS	F	Significance
Group	1	25.27	25.27	535.96	<.0001
Age	5	1.85	0.37	7.87	<.0001
Group by Age	5	0.35	0.07	1.48	0.1957
Error	588	27.72	0.05		
Total	599	55.19			

Be/Do (Be)

Source	df	SS	MS	F	Significance
Group	1	20.57	20.57	402.15	<.0001
Age	5	1.08	0.22	4.22	0.0009
Group by Age	5	0.36	0.07	1.41	0.2171
Error	588	30.07	0.05		
Total	599	52.08			

Be/Do (Do)

Source	df	SS	MS	F	Significance
Group	1	45.79	45.79	631.41	<.0001
Age	5	2.66	0.53	7.33	<.0001
Group by Age	5	0.5	0.1	1.39	0.2260
Error	588	42.64	0.07		
Total	599	91.59			

Elicited Grammar Composite

Source	df	SS	MS	F	Significance
Group	1	30.13	30.13	827.44	<.0001
Age	5	1.67	0.33	9.19	<.0001
Group by Age	5	0.21	0.04	1.17	0.3204
Error	588	21.41	0.04		
Total	599	53.43			

Table 4.17 t test for the Rice/Wexler Probe Scores and for the Elicited Grammar Composite

Probe Score/ Age Groups		Normal Group	Language Disorder Group	t	df	P
Third Person Singular						
3.00–3.05	M	0.71	0.29	4.98	61	<.0001
	SD	0.30	0.34			
3.06–3.11	M	0.80	0.26	7.31	72	<.0001
	SD	0.27	0.35			
4.00–4.05	M	0.87	0.38	8.15	98	<.0001
	SD	0.24	0.34			
4.06–4.11	M	0.91	0.39	9.60	98	<.0001
	SD	0.18	0.34			
5.00–5.05	M	0.93	0.47	8.63	98	<.0001
	SD	0.12	0.36			
5.06–5.11	M	0.97	0.47	10.12	98	<.0001
	SD	0.06	0.35			
6.00–6.05	M	0.97	0.57	7.66	98	<.0001
	SD	0.07	0.37			
6.06–6.11	M	0.96	0.57	7.82	98	<.0001
	SD	0.07	0.34			
Past Tense						
3.00–3.05	M	0.65	0.36	4.03	61	0.0002
	SD	0.25	0.31			
3.06–3.11	M	0.78	0.30	8.21	72	<.0001
	SD	0.22	0.25			
4.00–4.05	M	0.84	0.38	9.15	98	<.0001
	SD	0.19	0.30			
4.06–4.11	M	0.90	0.48	10.40	98	<.0001
	SD	0.10	0.27			
5.00–5.05	M	0.88	0.44	9.40	98	<.0001
	SD	0.12	0.31			
5.06–5.11	M	0.93	0.49	10.09	98	<.0001
	SD	0.08	0.30			
6.00–6.05	M	0.93	0.60	9.19	98	<.0001
	SD	0.08	0.24			
6.06–6.11	M	0.94	0.58	8.86	98	<.0001
	SD	0.06	0.28			
Be/Do (Be)						
3.00–3.05	M	0.72	0.23	5.92	61	<.0001
	SD	0.28	0.37			
3.06–3.11	M	0.86	0.40	6.04	72	<.0001
	SD	0.17	0.35			
4.00–4.05	M	0.87	0.48	7.45	98	<.0001
	SD	0.18	0.32			
4.06–4.11	M	0.90	0.57	6.31	98	<.0001
	SD	0.16	0.34			
5.00–5.05	M	0.93	0.46	10.03	98	<.0001
	SD	0.12	0.31			
5.06–5.11	M	0.93	0.60	8.76	98	<.0001
	SD	0.08	0.25			
6.00–6.05	M	0.96	0.59	9.18	98	<.0001
	SD	0.06	0.28			
6.06–6.11	M	0.96	0.62	8.39	98	<.0001
	SD	0.06	0.28			

Table 4.17 t test for the Rice/Wexler Probe Scores and for the Elicited Grammar Composite (continued)

Probe Score/ Age Groups		Normal Group	Language Disorder Group	t	df	P
Be/Do (Do)						
3.00–3.05	M	0.60	0.14	4.94	61	<.0001
	SD	0.35	0.33			
3.06–3.11	M	0.71	0.09	7.63	72	<.0001
	SD	0.36	0.24			
4.00–4.05	M	0.74	0.21	8.04	98	<.0001
	SD	0.34	0.32			
4.06–4.11	M	0.83	0.20	11.62	98	<.0001
	SD	0.23	0.31			
5.00–5.05	M	0.87	0.25	12.22	98	<.0001
	SD	0.18	0.31			
5.06–5.11	M	0.83	0.30	10.91	98	<.0001
	SD	0.20	0.28			
6.00–6.05	M	0.90	0.36	10.27	98	<.0001
	SD	0.13	0.35			
6.06–6.11	M	0.90	0.44	9.39	98	<.0001
	SD	0.14	0.32			
Elicited Grammar Composite						
3.00–3.05	M	0.67	0.25	7.13	61	<.0001
	SD	0.23	0.20			
3.06–3.11	M	0.79	0.26	11.45	72	<.0001
	SD	0.17	0.21			
4.00–4.05	M	0.83	0.36	10.46	98	<.0001
	SD	0.20	0.24			
4.06–4.11	M	0.89	0.41	12.97	98	<.0001
	SD	0.11	0.23			
5.00–5.05	M	0.90	0.41	12.55	98	<.0001
	SD	0.10	0.26			
5.06–5.11	M	0.92	0.47	12.33	98	<.0001
	SD	0.08	0.24			
6.00–6.05	M	0.94	0.53	11.85	98	<.0001
	SD	0.06	0.24			
6.06–6.11	M	0.94	0.55	10.78	98	<.0001
	SD	0.06	0.25			

3. Growth in tense-marking is not predicted by children's receptive vocabulary, nonverbal intelligence, or their mother's education (Rice, Wexler, & Hershberger, 1998).

   Longitudinal outcomes of the previous experimental study allowed for the calculation of the predictors of the children's observed growth, i.e., the changes in the children's tense-marking over time. Those calculations revealed that the change in grammatical tense performance over time was not predicted by the children's initial levels of receptive vocabulary, nonverbal intelligence, or their mother's education.

   Similar findings are evident in the standardization data, as reported in the validity section in Chapter 5. The association of the Elicited Grammar Composite is less with the more semantically loaded CELF subtests of Basic Concepts or Word Classes than with the more morphological subtests of Word Structure. The lower association is especially evident for the language disorder groups. This suggests that, although children can have both a grammatical marker and a semantic delay, the semantic performance is a poor predictor of the grammatical marker performance. Furthermore, for the language disorder group, the correlation between the grammatical marker and nonverbal IQ is very low, also suggesting that nonverbal IQ performance is a poor predictor of grammatical marker performance. Finally, performance on the Elicited Grammar Composite and Grammaticality Judgment probes is not predicted by parent education level for either the language disorder group or the normal language group. It is not accurate to assume that children with low levels of performance on the Rice/Wexler are likely to have parents with low levels of education.

4. The grammatical marker is not evident in all areas of morphology (Rice, Wexler, & Cleave, 1995; Rice & Wexler, 1996; Rice, Wexler, & Hershberger, 1998).

   Earlier studies show that at the same time the language disorder group's performance is low on the grammatical marker relative to control groups, their performance on morphemes that do not mark finiteness is at levels similar to that of controls. This is evident, for example, in the children's use of regular plural *-s* affix, which is at high levels of accuracy at the same time that the phonologically similar third person singular *-s* is at low levels of accuracy. This suggests further that affected children can know important elements of morphology, but the particular area of finiteness marking can be difficult for them. For this reason, the test focuses on finiteness marking and not all areas of morphology.

   Another unaffected morpheme is that of progressive *-ing*, as in "Patsy is talking." At the same time that children with language impairments are likely to omit copula or auxiliary forms of *Be*, they are unlikely to omit the *-ing* affix. This is important because the areas of morphological strength, such as regular plurals and progressive *-ing*, can be used in a grammaticality judgment task to further evaluate their grammatical understanding.

5. A. Children's judgments of simple clauses show outcomes very similar to their productions, i.e., affected children, as a group, are likely to accept the kinds of utterances they generate (i.e., with omitted finiteness markers), whereas control children are more likely to judge these utterances to be poorly formed (Rice, Wexler, & Redmond, 1999).

B. The growth in grammaticality judgment accuracy over time shows a similar trajectory as growth in grammatical productions, and the predictors of growth are similar, as well (Rice, Wexler & Redmond, 1999).

The earlier research studies established the important finding that the grammaticality judgment tasks show the expected parallels with the children's production data, which in turn is a crucial indicator of validity of measurement of the targeted grammatical property.

Detailed analyses of the test data provide further evidence of the principled growth over time in children's judgments of the grammar marker, and the fact that children are less accurate at judging grammatical violations with dropped finiteness markers than at judging dropped *-ing* or subject-verb agreement violations.

For those readers following the literature closely, note that the labels for the A' values for the test have been changed from the labels in the earlier research reports: "Optional Infinitive A' (OI)" is now "Dropped Marker A'"; "Bad Agreement A' (BA)" is now "Agreement A'"; and "Drop -ing A'" is now "Dropped-ing A'."

An ANOVA of group (language disorder versus normal) by age level, by judgment type shows that there are significant differences between the two groups, there are significant changes over age, and there are significant differences among the three kinds of A' judgment values. There are no significant interactions, indicating that the performance of the language disorder group is parallel to that of the normal group, but at lower levels of performance. Table 4.18 reports the results of this analysis.

Table 4.18 Repeated Measures ANOVA—Group by Age and Judgment Type

Source	df	SS	MS	F	p
Group	1	33.48	33.48	228.05	<.0001
Age	5	11.58	2.32	15.78	<.0001
Group by Age	5	0.64	0.13	0.87	0.4998
A' Judgment	2	1.33	0.66	39.91	<.0001
Group by A'	2	0.08	0.04	2.22	0.1091
Age by A'	10	0.22	0.02	1.30	0.2271
Group by Age by A'	10	0.12	0.01	0.69	0.7314
Error	1176	20.05	0.02		

To better reveal the differences within the A' measures for the language disorder groups and the normal groups, an ANOVA was carried out separately for the language disorder groups and the normal groups. The outcomes were the same for the language disorder groups as for the normal language groups: a significant age group effect showing that older children performed better than younger children, Dropped Marker A' is less accurate than Agreement A' and less accurate than Dropped -ing A', and Dropped -ing A' is not different from Agreement A'. These data are reported in Table 4.19.

Table 4.19 Repeated Measures ANOVA Table—Normal Language Group and Language Disorder Group

Repeated Measures ANOVA—Normal Language Group

Source	df	SS	MS	F	Sig
Age	5	3.9	0.78	13.16	<.0001
Dropped Marker and Agreement	1	0.27	0.27	41.58	<.0001
Error	294	1.88	0.01		
Age	5	3.93	0.79	11.96	<.0001
Dropped Marker and Dropped -ing	1	0.31	0.31	31.89	<.0001
Error	294	2.85	0.01		
Age	5	3.61	0.72	10.47	<.0001
Agreement and Dropped -ing	1	0.002	0.002	0.32	0.5694
Error	294	1.45	0.005		

Repeated Measures ANOVA—Language Disorder Group

Source	df	SS	MS	F	Sig
Age	7	14.68	2.1	19.01	<.0001
Dropped Marker and Agreement	1	0.82	0.82	58.89	<.0001
Error	392	5.45	0.01		
Age	7	15.42	2.2	18.44	<.0001
Dropped Marker and Dropped -ing	1	1.17	1.17	40.62	<.0001
Error	392	11.27	0.03		
Age	7	15.63	2.23	16.47	<.0001
Agreement and Dropped -ing	1	0.03	0.03	1.24	0.2654
Error	392	9.77	0.02		

Follow-up *t*-test analysis looked at whether or not the group differences were apparent at each age level for each of the A' indices. Those are reported in Tables 4.20 and 4.21. It can be seen that the normal group performed significantly better than the language disorder group at each age level for each of the three A' indices, at probability levels of .0001–.0003.

Table 4.20 *t* test Between Rice/Wexler Grammaticality Judgment Probe Scores

Probe Score/ Age		Normal Group	Language Disorder Group	t	df	P
Dropped Marker						
4.00–4.05	M	0.70	0.43	4.84	98	<.0001
	SD	0.25	0.30			
4.06–4.11	M	0.75	0.43	6.33	98	<.0001
	SD	0.21	0.29			
5.00–5.05	M	0.80	0.53	6.35	98	<.0001
	SD	0.20	0.22			
5.06–5.11	M	0.83	0.58	6.69	98	<.0001
	SD	0.18	0.20			
6.00–6.05	M	0.92	0.58	8.21	98	<.0001
	SD	0.12	0.26			
6.06–6.11	M	0.93	0.63	8.61	98	<.0001
	SD	0.09	0.22			
Agreement						
4.00–4.05	M	0.75	0.46	4.68	98	<.0001
	SD	0.25	0.35			
4.06–4.11	M	0.81	0.50	5.59	98	<.0001
	SD	0.20	0.34			
5.00–5.05	M	0.84	0.60	5.20	98	<.0001
	SD	0.21	0.25			
5.06–5.11	M	0.87	0.65	5.18	98	<.0001
	SD	0.17	0.25			
6.00–6.05	M	0.94	0.65	6.70	98	<.0001
	SD	0.12	0.29			
6.06–6.11	M	0.98	0.73	8.26	98	<.0001
	SD	0.05	0.20			
Dropped -ing						
4.00–4.05	M	0.73	0.45	4.05	98	0.0001
	SD	0.31	0.38			
4.06–4.11	M	0.83	0.53	4.90	98	<.0001
	SD	0.22	0.38			
5.00–5.05	M	0.85	0.64	3.93	98	0.0002
	SD	0.24	0.30			
5.06–5.11	M	0.87	0.69	3.72	98	0.0003
	SD	0.17	0.29			
6.00–6.05	M	0.94	0.61	6.37	98	<.0001
	SD	0.15	0.34			
6.06–6.11	M	0.97	0.71	6.51	98	<.0001
	SD	0.07	0.27			

A follow-up series of analyses examined whether or not the observed higher levels of Agreement A' compared to Dropped Marker A' were statistically significant at each age level for the language disorder group. A series of *t* tests found that the differences were statistically significant for each age level, except for the youngest age group where the obtained probability was .16. Table 4.21 reports the means, differences, t value and probability between each of the A' values.

Table 4.21 *t* test Between Rice/Wexler Grammaticality Judgment Probe Scores—Children in the Language Disorder Group

Difference Between Dropped Marker and Agreement

Age Group		Dropped Marker	Agreement	t	df	p
4.00–4.05	M	0.43	0.46	]1.42	49	0.1609
	SD	0.30	0.35			
4.06–4.11	M	0.43	0.50	]2.43	49	0.0189
	SD	0.29	0.34			
5.00–5.05	M	0.53	0.60	]3.10	49	0.0032
	SD	0.22	0.25			
5.06–5.11	M	0.58	0.65	]2.69	49	0.0098
	SD	0.20	0.25			
6.00–6.05	M	0.58	0.65	]2.88	49	0.0059
	SD	0.26	0.29			
6.06–6.11	M	0.63	0.73	]3.44	49	0.0012
	SD	0.23	0.20			

Difference Between Dropped Marker and Dropped -ing

Age Group		Dropped Marker	Dropped -ing	t	df	p
4.00–4.05	M	0.43	0.45	]0.62	49	0.5392
	SD	0.30	0.38			
4.06–4.11	M	0.43	0.52	]2.11	49	0.0403
	SD	0.30	0.38			
5.00–5.05	M	0.53	0.64	]3.44	49	0.0012
	SD	0.22	0.30			
5.06–5.11	M	0.58	0.69	]3.37	49	0.0015
	SD	0.20	0.29			
6.00–6.05	M	0.58	0.61	]0.57	49	0.5723
	SD	0.26	0.34			
6.06–6.11	M	0.63	0.71	]2.46	49	0.0176
	SD	0.23	0.27			

Difference Between Agreement and Dropped -ing

Age Group		Agreement	Dropped -ing	t	df	p
4.00–4.05	M	0.46	0.45	0.27	49	0.7892
	SD	0.35	0.38			
4.06–4.11	M	0.50	0.53	]0.62	49	0.5398
	SD	0.34	0.38			
5.00–5.05	M	0.60	0.64	]1.08	49	0.2867
	SD	0.25	0.30			
5.06–5.11	M	0.65	0.69	]1.15	49	0.2546
	SD	0.25	0.29			
6.00–6.05	M	0.65	0.61	1.08	49	0.2842
	SD	0.29	0.34			
6.06–6.11	M	0.73	0.71	0.61	49	0.5448
	SD	0.20	0.27			

These outcomes replicate the patterns reported in the earlier research studies. Although the overall level of performance of the language disorder groups is markedly below that of the normal groups, at the same time, the language disorder groups are more likely to do even more poorly on grammaticality judgments that parallel their production data (i.e., simple sentences in which finiteness markers are omitted). The new evidence is that the youngest group of affected children for whom there are grammaticality judgment data, ages 4.00–4.05, may be at the beginning levels of being able to make these judgments or understand the task. As noted in Chapter 3, the clinical implication is that for individual children it is important to consider if their A' values are low at all three A' indicators, or if they perform lower on the Dropped Marker A'. Low performance across all three indicators may indicate a generally more immature ability to evaluate grammatical forms. Conversely, if there is better performance on Dropped *-ing*, for example, then intervention activities beginning with this morpheme for teaching grammaticality judgments may be appropriate, later moving to judgments of sentences with the dropped grammatical marker.

6. Growth in regular and irregular past tense verb markings follows different trajectories, and is associated with different predictor variables. At the same time, the underlying knowledge of finiteness marking unifies the two different ways to express past tense (Rice, Wexler, Marquis, & Hershberger, 2000).

   The composite measure of past tense on the test is that of finiteness marking, which credits a child with an overregularized form of regular past tense on an irregular verb form, such as "*catched" for "caught." This way of combining performance on the two classes of verbs was thoroughly evaluated in the earlier experimental studies and reported in detail by Rice, Wexler, Marquis, & Hershberger (2000).

   The outcomes of the test, with a slightly modified form of the elicitation probe, also clearly reveal differences in the growth over time in the children's performance on the regular and irregular past tense items, and a greater similarity between regular past tense performance levels and the composite finite past tense calculation than between irregular past tense performance levels and the composite finite past tense calculation.

   As shown in Figures 4.11 and 4.12, these outcomes make it very clear that different acquisition mechanisms are at work for the different forms of verbs. At the same time, they share an underlying property; the likelihood of attaching a past tense marker, be it the regular *-ed* affix or one of the irregular morphemes.

7. Children with language impairment who also have nonverbal IQ levels somewhat below normal range perform at lower levels on the grammar marker than do children with Specific Language Impairment (Rice & Tomblin, 1999).

   On the experimental version of the elicitation and judgment probes, the performance of a large sample of children with SLI and children with both language impairments and borderline nonverbal IQ levels was compared to that of normal groups in the 5.00–8.11 age range. Throughout this age range, the children with lower IQ levels performed below the SLI group and below the normal controls. This suggests that difficulties with the grammatical marker may be a "tip of the iceberg," and further suggests that the identification of difficulties in this area may serve to detect children with low levels of language acquisition relative to children of the same age.

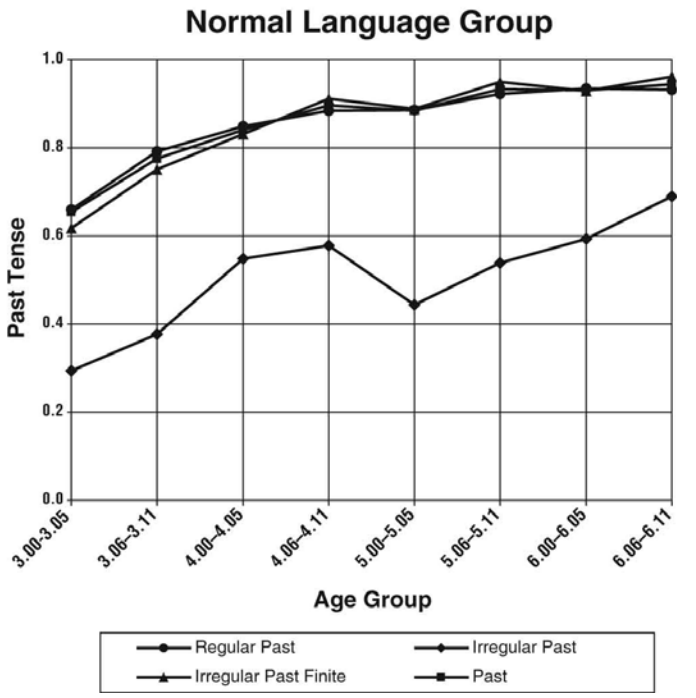


Figure 4.11 Past Tense Calculations—Normal Language Group

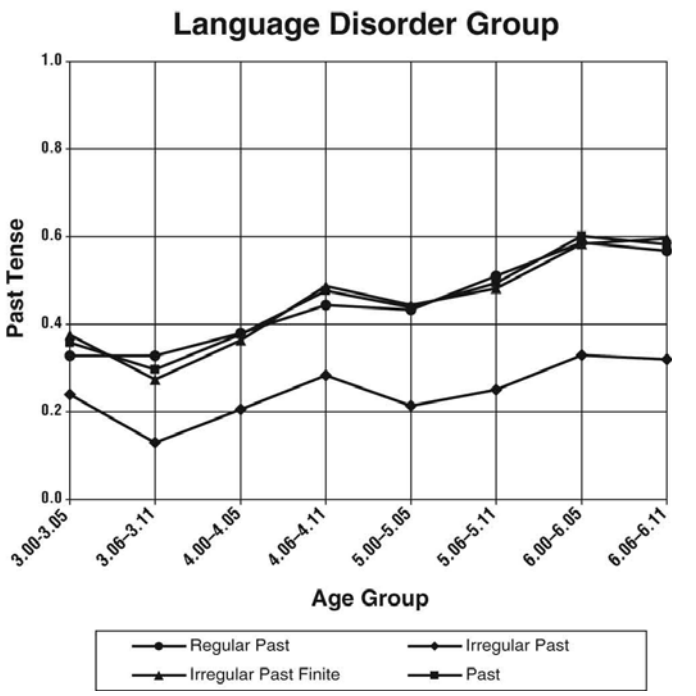


Figure 4.12 Past Tense Calculations—Language Disorder Group

Readers interested in further evidence relevant to the grammatical marker in diverse populations of children are encouraged to consult the following sources: 1) Rice, Spitz, & O'Brien (1999) found that 4-year-old children with a history of time in the neonatal intensive care unit at birth performed lower as a group on past tense morphology than did children with a normal birth history. 2) Rice, Mervis, Klein, & Rice (1999) found that children with Williams syndrome performed higher on finiteness marking than did children with SLI at the same levels of mean length of utterance. 3) Bedore & Leonard (1998) report that tense-marking is lower in children with SLI than control children. 4) Norbury, Bishop, & Briscoe (2001) report that tense-marking is lower than control groups for children with SLI and children with hearing impairment. 5) Oetting & McDonald (2001) report that children with non-mainstream dialect use who have language impairments perform lower in their use of irregular past tense than dialect speakers without language impairments. 6) Paradis, Crago, Genesee, & Rice (2001) found that bilingual French-English-speaking children with SLI were less accurate in tense morphology in French and in English. 7) Hansson, Nettelbladt & Leonard (2000) report that Swedish-speaking children with SLI perform lower than control children on tense-marking morphology.

Technical Characteristics 5

This chapter presents information about the technical properties of the Rice/Wexler Test of Early Grammatical Impairment; specifically, the traditional aspects of reliability and validity.

Reliability

A test is reliable to the degree that scores are consistent over repeated testing administrations. Reliability must be estimated because it is not directly observable. The reliability of the Rice/Wexler was estimated by using its test-retest reliability.

Test-Retest Reliability

One way of estimating the stability of an instrument is to examine its test-retest reliability. The stability of an instrument is measured by administering the instrument to one group of subjects on two separate occasions and comparing the scores. For this to be a meaningful estimate of reliability, the trait being measured must be stable, the test must not produce large practice effects, and the subjects must not change significantly on the trait between administrations of the test. In practice, this means that the time between test administrations should be long enough to minimize such effects but not so long that subjects change in terms of the construct being measured. If these conditions are not met, the test-retest correlation may not provide a meaningful estimate of the reliability of the test.

The *Rice/Wexler Test of Early Grammatical Impairment* test-retest reliability was evaluated based on the performance of 106 children who were part of the standardization research study. The sample consisted of 54 children who were between the ages of 4.00 and 4.05, and 52 children who were between the ages of 4.06 and 4.11. The sample was 55% female and 45% male. The racial/ethnic representation was 22% African American, 6% Hispanic, 70% White, and 2% other racial/ethnic origins. The children took the test on two separate occasions, administered by the same examiner. The children in the sample took the *Rice/Wexler Test of Early Grammatical Impairment* and, depending on the age of the child, two subtests from the *CELF®–Preschool* (Word Structure and Basic Concepts) or *CELF®–3* (Word Structure and Word Classes) during the first administration, and then took only the *Rice/Wexler Test of Early Grammatical Impairment* during the second administration.

The time interval between administrations of the *Rice/Wexler Test of Early Grammatical Impairment* was 7 to 21 days. Table 5.1 presents the means and standard deviations for both administrations, and the stability coefficients and mean

absolute score differences between the test and retest by area for all subjects combined. The test-retest correlation coefficients are provided for the five areas of the test (four core probe scores and the Elicited Grammar Composite) and for the Grammaticality Judgment Probe. The retest values for the core probes and the Elicited Grammar Composite ranged from .82 to .95. The highest correlation occurred in the Elicited Grammar Composite. The lowest correlation occurred on the Past Tense Probe. For the Grammaticality Judgment Probe the retest values were different for the younger children, ages 4.00–4.05, than for the older children, ages 4.06–4.11. For the younger children, which included children performing near chance levels, retest values ranged from .37 to .44. For this group the highest correlation occurred on Agreement. The values for Dropped Marker and Dropped *-ing* for these ages were the same. For the older group, whose performance levels are higher, the retest values ranged from .65 to .82. For this group, the highest correlation occurred on the Dropped Marker. The lowest correlation occurred on the Dropped *-ing*.

To further evaluate the stability of the Rice/Wexler, the mean absolute score differences between the test and retest were examined. Mean absolute score differences enable you to see the amount of variation in scores, regardless of whether the difference was positive or negative. The mean absolute score differences on the Rice/Wexler were small, ranging from .055 to .092. The high correlations obtained on the retest study and the small mean absolute score differences support adequate test-retest stability for the Rice/Wexler.

Table 5.1 Test-Retest Means, Standard Deviations, Mean Absolute Score Differences, and Stability Coefficients for Rice/Wexler Probes and Elicited Grammar Composite (n = 106)

	1st Administration		2nd Administration			
Probe	**Mean**	**SD**	**Mean**	**SD**	**diff**	**r**
Third Person Singular	.79	.33	.80	.32	.07	.92
Past Tense	.81	.23	.83	.25	.08	.82
Be/Do (Be)	.84	.21	.84	.21	.07	.87
Be/Do (Do)	.69	.34	.71	.32	.09	.88
Elicited Grammar Composite	.78	.24	.80	.24	.06	.95
Grammaticality Judgment						
(Ages 4.00–4.05)						
Dropped Marker	.61	.27	.67	.27	.19	.37
Agreement	.67	.27	.72	.26	.17	.43
Dropped -ing	.69	.31	.74	.31	.20	.37
Grammaticality Judgment						
(Ages 4.06–4.11)						
Dropped Marker	.71	.24	.73	.23	.09	.82
Agreement	.77	.25	.78	.22	.10	.80
Dropped -ing	.75	.32	.80	.24	.14	.65

Validity

Validity studies examine the extent to which evidence and theory support specific interpretation of the test scores. "It is the interpretations of test scores required by proposed uses that are evaluated, not the test itself" (The Standards for Education and Psychological Testing, 1999). The sources of evidence depend on the intended interpretations for the test scores. Important validity evidence can be obtained by examining the relationship between the test content and the construct it is intended to measure (evidence based on test content) and relationships among subtests on which the test score interpretations are based (evidence based on internal structure).

The validity evidence of an instrument can be accumulated through multiple studies using different methods. This process was begun during research investigations into the theoretical applicability of "Optional Infinitives" and the identification of Specific Language Impairment (SLI), by these authors and other researchers in the field, and during the development of the *Rice/Wexler Test of Early Grammatical Impairment*. Research in this area of child language is ongoing by these authors and by other researchers.

Evidence Based on Test Content

Validity evidence related to test content is supported when the content area(s) being measured are generally accepted as the proposed construct (content relevance) and when the content areas are accepted to be an adequate sampling of these areas (content coverage).

Both of these evidences of content-related validity are well supported for the *Rice/Wexler Test of Early Grammatical Impairment*. Much research has been conducted and published recently in the area of Specific Language Impairment; more specifically, a significant amount of research has been focused on investigating a grammatical marker. Information provided in the Purpose and Theoretical Background sections of Chapter 1 and in the History of the Instrument in Chapter 4 describes the theoretical foundation upon which this test was based and the rationale for the unique make-up of each probe. See the reference list for published research documenting this set of clinical markers.

Evidence Based on Internal Structure

Validity evidence based on the internal structure of a test can indicate the degree to which the test items and subtests or sections relate to each other in predictable ways, according to the theory or construct on which the test is based. The internal structure of the Rice/Wexler was evaluated by exploring the relationship between areas of the test to determine if predicted results (supporting validity) would be observed.

Each probe score and the Elicited Grammar Score was evaluated in terms of how they relate to each other. It would be expected that there would be moderate to low correlations between each of these scores. As discussed in Chapter 3, although each probe is used to measure an element of grammar, each probe addresses a different specific element of grammar. The morphemes share the property of finiteness and differ in other ways, such as whether the morphemes are affixes on lexical verbs versus the free-standing morphemes of *Be* and *Do*, and whether the morpheme appears with progressive verbs (i.e., *Be* auxiliary) or predicate adjectives (i.e., *Be* copula), and whether the morphemes can move to the front of the sentence to form questions (such as *Be* and *Do* forms) or if they cannot move to the front for questions (past tense and third person singular present tense).

Table 5.2 Correlations of the Rice/Wexler probes and Elicited Grammar Composite—Children in the Normal Language Group

Age	Subtest	Third Person Singular	Past Tense	Be/Do (Be)	Be/Do (Do)	Elicited Grammar Composite
3.00–3.05 (n = 43)	Third Person Singular		.22	.45	.31	.66
	Past Tense	.22		.44	.52	.70
	Be/Do (Be)	.45	.44		.62	.83
	Be/Do (Do)	.31	.52	.62		.83
	Elicited Grammar Composite	.66	.70	.83	.83	
3.06–3.11 (n = 50)	Third Person Singular		.42	.03	.11	.61
	Past Tense	.42		.29	.21	.68
	Be/Do (Be)	.03	.29		.39	.57
	Be/Do (Do)	.11	.21	.39		.74
	Elicited Grammar Composite	.61	.68	.57	.74	
4.00–4.05 (n = 50)	Third Person Singular		.74	.37	.63	.83
	Past Tense	.74		.51	.57	.82
	Be/Do (Be)	.37	.51		.64	.73
	Be/Do (Do)	.63	.57	.64		.90
	Elicited Grammar Composite	.83	.82	.73	.90	
4.06–4.11 (n = 50)	Third Person Singular		.15	.21	.08	.52
	Past Tense	.15		.35	.40	.59
	Be/Do (Be)	.21	.35		.53	.77
	Be/Do (Do)	.08	.40	.53		.81
	Elicited Grammar Composite	.52	.59	.77	.81	
5.00–5.05 (n = 50)	Third Person Singular		.63	.30	.23	.65
	Past Tense	.63		.50	.51	.82
	Be/Do (Be)	.30	.50		.66	.80
	Be/Do (Do)	.23	.51	.66		.83
	Elicited Grammar Composite	.65	.82	.80	.83	
5.06–5.11 (n = 50)	Third Person Singular		.41	.16	.19	.45
	Past Tense	.41		.38	.37	.66
	Be/Do (Be)	.16	.38		.64	.78
	Be/Do (Do)	.19	.37	.64		.90
	Elicited Grammar Composite	.45	.66	.78	.90	
6.00–6.05 (n = 50)	Third Person Singular		.12	.07	.13	.48
	Past Tense	.12		.24	.06	.52
	Be/Do (Be)	.07	.24		.46	.65
	Be/Do (Do)	.13	.06	.46		.79
	Elicited Grammar Composite	.48	.52	.65	.79	
6.06–6.11 (n = 50)	Third Person Singular		.26	.18	.16	.52
	Past Tense	.26		.30	.28	.58
	Be/Do (Be)	.18	.30		.51	.69
	Be/Do (Do)	.16	.28	.51		.85
	Elicited Grammar Composite	.52	.58	.69	.85	

Table 5.3 Correlations of the Rice/Wexler probes and Elicited Grammar Composite—Children With Language Impairments

Age	Subtest	Third Person Singular	Past Tense	Be/Do (Be)	Be/Do (Do)	Elicited Grammar Composite
3.00–3.05 (n = 20)	Third Person Singular		.17	].31	].27	.22
	Past Tense	.17		.35	.12	.66
	Be/Do (Be)	].31	.35		.82	.80
	Be/Do (Do)	].27	.12	.82		.72
	Elicited Grammar Composite	.22	.66	.80	.72	
3.06–3.11 (n = 24)	Third Person Singular		.65	.56	.06	.85
	Past Tense	.65		.36	.02	.72
	Be/Do (Be)	.56	.36		.16	.80
	Be/Do (Do)	.06	.02	.16		.38
	Elicited Grammar Composite	.85	.72	.80	.38	
4.00–4.05 (n = 50)	Third Person Singular		.52	.51	.40	.81
	Past Tense	.52		.31	.40	.72
	Be/Do (Be)	.51	.31		.49	.76
	Be/Do (Do)	.40	.40	.49		.75
	Elicited Grammar Composite	.81	.72	.76	.75	
4.06–4.11 (n = 50)	Third Person Singular		.40	.46	.52	.82
	Past Tense	.40		.35	.18	.62
	Be/Do (Be)	.46	.35		.42	.77
	Be/Do (Do)	.52	.18	.42		.73
	Elicited Grammar Composite	.82	.62	.77	.73	
5.00–5.05 (n = 50)	Third Person Singular		.51	.66	.39	.81
	Past Tense	.51		.55	.55	.80
	Be/Do (Be)	.66	.55		.57	.86
	Be/Do (Do)	.39	.55	.57		.76
	Elicited Grammar Composite	.81	.80	.86	.76	
5.06–5.11 (n = 50)	Third Person Singular		.68	.62	.53	.87
	Past Tense	.68		.57	.57	.85
	Be/Do (Be)	.62	.57		.57	.81
	Be/Do (Do)	.53	.57	.57		.79
	Elicited Grammar Composite	.87	.85	.81	.79	
6.00–6.05 (n = 50)	Third Person Singular		.40	.63	.37	.80
	Past Tense	.40		.39	.56	.72
	Be/Do (Be)	.63	.39		.43	.79
	Be/Do (Do)	.37	.56	.43		.77
	Elicited Grammar Composite	.80	.72	.79	.77	
6.06–6.11 (n = 50)	Third Person Singular		.55	.57	.47	.81
	Past Tense	.55		.55	.61	.83
	Be/Do (Be)	.57	.55		.47	.79
	Be/Do (Do)	.47	.61	.47		.79
	Elicited Grammar Composite	.81	.83	.79	.79	
7.00–7.11 (n = 50)	Third Person Singular		.39	.60	.44	.84
	Past Tense	.39		.25	.40	.65
	Be/Do (Be)	.60	.25		.43	.74
	Be/Do (Do)	.44	.40	.43		.77
	Elicited Grammar Composite	.84	.65	.74	.77	
8.00–8.11 (n = 50)	Third Person Singular		.54	.55	.45	.82
	Past Tense	.54		.55	.52	.79
	Be/Do (Be)	.55	.55		.77	.85
	Be/Do (Do)	.45	.52	.77		.81
	Elicited Grammar Composite	.82	.79	.85	.81	

Moderate to low correlations would represent the shared properties among the set of morphemes, the shared function of finiteness-marking. Table 5.2 (page 88) provides the correlation coefficients between the Third Person Singular Probe, the Past Tense Probe, the *Be* score and the *Do* score from the Be/Do Probe and the Elicited Grammar Composite for children in the normal language group. Table 5.3 (page 89) provides this information for children with language impairments. For the children in the normal language group, the correlations are in the expected moderate to low range. When interpreting these values, keep in mind that the generally high levels of performance for the children in the normal language group, and the resultant restricted variance, can affect the calculations of correlation in ways that lead to lower values. This may be occurring at the oldest ages of the normal groups. For children with language impairments (with the exception of the lowest ages), the correlations among the individual probes are generally in the moderate range (.40–.60) or above. When interpreting these values, keep in mind the possible effect of uniformly low performance within the group i.e., floor effects, which, like ceiling effects can affect the calculations in ways that lead to lower values. This may contribute to the outcomes for the younger children where floor effects are more evident.

Evidence Based on Relations to Other Variables

An important type of validity evidence is obtained by comparing the scores obtained from the instrument of interest with the scores obtained on other variables. When these variables are well-established or accepted instruments that purport to measure the same thing, high correlations provide evidence of convergent validity. In addition, it is often important to show that the test correlates less highly with related, but different constructs. This provides evidence of discriminant validity.

If the Rice/Wexler is, in fact, a test of grammar or grammatical skills, one would expect that an individual's performance would more highly correlate with another measure of grammar skills than with a measure of another area of language. However, because the Rice/Wexler is unique in format and theoretical construct, no existing instruments are available or appropriate for direct comparison.

Therefore, in order to provide evidence of convergent and discriminant validity for the Rice/Wexler, a study was conducted to evaluate the relationship between performance on the *Rice/Wexler Test of Early Grammatical Impairment* and selected subtests of *CELF®–3* or *CELF®–Preschool*. Because there was overlap in the ages in the CELF tests, it was decided to set the age groups so that a specific age group of children would take each subtest. The selected subtests from each of these tests include one subtest documented in its test manual as testing a similar area of language as the Rice/Wexler and one subtest documented as testing a different area of language than the Rice/Wexler.

All children in the standardization research study between the ages of 3.00 and 5.11 completed the *Rice/Wexler Test of Early Grammatical Impairment* and the Word Structure and Basic Concepts subtests of the *CELF®–Preschool*. All children in the standardization research study between the age of 6.00 and 8.11 completed the *Rice/Wexler Test of Early Grammatical Impairment* and the Word Structure and Word Classes subtests of the *CELF®–3*.

Convergent Evidence

Convergent validity evidence is examined by investigating the relationship between scores from the test under investigation and another measure that is purported to be evaluating or measuring a similar construct.

If the Rice/Wexler measures specific aspects of grammar, one would expect scores on the test to correlate with other measures of general grammar skills. To evaluate this, the probe scores and the Elicited Grammar Composite from the Rice/Wexler were compared to the standard scores from Word Structure subtests of *CELF®–Preschool* and *CELF®–3*. According to the test manuals, the Word Structure subtest from each of these instruments measures a child's knowledge and use of morphological rules. The general content may then be considered similar to the content measured in the Rice/Wexler. However, one significant distinction between the two instruments is that in the CELF instruments the content of the Word Structure subtest covers a larger set of morphological structures than are included in the Rice/Wexler. The reasons for this smaller set in the Rice/Wexler lie in the theoretical basis of the instrument; that only a few morphological structures constitute the clinical marker. Therefore one would expect moderate correlations between these measures.

Relationship Between the *Rice/Wexler Test of Early Grammatical Impairment* and Word Structure From *CELF®–Preschool*

This study included a total of 265 children between the ages of 3.00–5.11 who were diagnosed with a language impairment and 435 children between the ages of 3.00–5.11 who were in the normal language group. In the group of children diagnosed with a language impairment, there were 68% males and 32% females; the racial/ethnic representation was 78% White, 9% African American, 8% Hispanic, and 5% of other races. In the normal language group, there were 48% males and 52% females; the racial/ethnic representation was 64% White, 16% African American, 16% Hispanic, and 4% of other races. For both age groups (3.00–5.11 and 6.00–8.11), the two tests were administered by the same examiner immediately after or within one week of the administration of the Rice/Wexler. This group of children (normal language group and children with language impairments) was included in both studies examining the relationship between Rice/Wexler and subtests from *CELF®–Preschool*.

Table 5.4 (page 92) reports the means, standard deviations, and correlations between *CELF®–Preschool* Word Structure and Rice/Wexler probes and Elicited Grammar Composite for children ages 3.00–5.11.

As expected, moderate correlations are observed between the subtests with the probe outcomes, ranging from .40 to .53 for the children in the normal language group, and .32 to .47 for the children with language impairments.

Table 5.4 Means, Standard Deviations, and Correlations Between *CELF®-Preschool* Word Structure and Rice/Wexler Probes and Elicited Grammar Composite

Subtest		Normal Language Group Ages 3.00–5.11 (n = 435)	Language Disorder Group Ages 3.00–5.11 (n = 265)
Word Structure	Mean	11.43	6.66
	SD	3.09	2.88
Elicited Grammar Composite	Mean	.77	.40
	SD	.25	.26
	r	.53	.47
Third Person Singular	Mean	.78	.42
	SD	.32	.36
	r	.40	.43
Past Tense	Mean	.78	.43
	SD	.25	.30
	r	.45	.39
Be/Do (Be)	Mean	.81	.51
	SD	.25	.33
	r	.47	.34
Be/Do (Do)	Mean	.71	.24
	SD	.34	.33
	r	.48	.32

Relationship Between the Rice/Wexler Test of Early Grammatical Impairment and Word Structure From *CELF®–3*

This study included a total of 279 children between the ages of 6.00–8.11 who were diagnosed with a language impairment and 149 children between the ages of 6.00–8.11 who were in the normal language group. In the group of children diagnosed with a language impairment, there were 63% males and 37% females; the racial/ethnic representation was 65% White, 18% African American, 7% Hispanic, and 10% of other races. In the normal language group, there were 47% males and 53% females; the racial/ethnic representation was 74% White, 5% African American, 8% Hispanic, and 13% of other races. For both age groups, the two tests were administered by the same examiner immediately after or within one week of the administration of the Rice/Wexler. This group of children (normal language group and children with language impairments) was included in both studies examining the relationship between Rice/Wexler and subtests from *CELF®–3*.

Table 5.5 reports the means, standard deviations, and correlations between *CELF®–3* Word Structure and Rice/Wexler probes and Elicited Grammar Composite for children ages 6.00–8.11.

Table 5.5 Means, Standard Deviations, and Correlations Between *CELF®–3* Word Structure and Rice/Wexler Probes and Elicited Grammar Composite

Subtest		Normal Language Group Ages 6.00–6.11 (n = 149)	Language Disorder Group Ages 6.00–8.11 (n = 279)
Word Structure	Mean	12.20	7.28
	SD	2.87	2.91
Elicited Grammar Composite	Mean	.91	.67
	SD	.12	.25
	r	.53	.54
Third Person Singular	Mean	.92	.68
	SD	.18	.35
	r	.40	.41
Past Tense	Mean	.91	.69
	SD	.14	.26
	r	.37	.42
Be/Do (Be)	Mean	.94	.72
	SD	.10	.26
	r	.45	.45
Be/Do (Do)	Mean	.87	.57
	SD	.16	.33
	r	.48	.48

Again, there are moderate correlations ranging from .37 to .53 for the children in the normal language group and .41 to .54 for the children with language impairments. Overall, there is consistent evidence of a moderate correlation association between the Word Structure subtests from the CELF tests and the probes from the Rice/Wexler. These results indicate that there is some shared component between these two measures. It should be noted that the same pattern is seen among both the normal and language disordered populations. Although we have evidence of the moderate correlations that would be expected, it is important to develop discriminant evidence that other, less related criteria yield even lower correlations. This is discussed in the following section.

Discriminant Evidence

Discriminant evidence of validity is provided by investigating or studying the relationship between scores from the test under investigation and another measure that is purported to be evaluating or measuring a different, or only peripherally related, construct. If such correlations are lower than the ones developed for convergent validity, such a pattern helps validate the test.

If the Rice/Wexler measures aspects of grammar, you would expect to find a very low correlation between scores on that test and scores on another measure that does not purport to measure grammar skills, or that, in fact, purports to test something else. To evaluate this, probe scores and the Elicited Grammar Composite from the Rice/Wexler were compared to the standard scores from the Basic Concepts subtest of *CELF®–Preschool* (for children ages 3.00–5.11) and the Word Classes subtest of *CELF®–3* (for children ages 6.00–8.11).

The Basic Concepts subtest, according to the *CELF®–Preschool* test manual, measures a child's knowledge of modifiers, including attributes, number/quantity, dimension/size, direction/location/position, and equality. The Word Classes subtest of *CELF®–3* evaluates the associative relationships between words. The general content of both of these subtests would be considered different from the content measured in the Rice/Wexler. It should be noted that although the specific content is different, all three instruments measure aspects of language, therefore some degree of correlation would be expected. However, the overall correlation expected would be very low, as the predominant nature of the content differs between these measures.

Relationship Between the Rice/Wexler Test of Early Grammatical Impairment and Basic Concepts Subtest From *CELF®–Preschool*

Table 5.6 reports the means, standard deviations, and correlations between *CELF®–Preschool* Basic Concepts and Rice/Wexler probes and Elicited Grammar Composite for children ages 3.00–5.11.

The correlations for the children in the normal language group range from .31 to .40. As expected these are lower than the earlier convergent correlations which ranged from .40 to .53. For the children with language impairments, the range is .12 to .21, which again is lower than the convergent correlations of .32 to .47.

Table 5.6 Means, Standard Deviations, and Correlations Between *CELF®–Preschool* Basic Concepts and Rice/Wexler probes and Elicited Grammar Composite

Subtest		Normal Language Group Ages 3.00–5.11 (n = 435)	Language Disorder Group Ages 3.00–5.11 (n = 265)
Basic Concepts	Mean	11.33	7.60
	SD	3.08	3.33
Elicited Grammar Composite	Mean	.77	.40
	SD	.25	.26
	r	.40	.21
Third Person Singular	Mean	.78	.42
	SD	.32	.36
	r	.32	.18
Past Tense	Mean	.78	.43
	SD	.25	.30
	r	.31	.18
Be/Do (Be)	Mean	.81	.51
	SD	.25	.33
	r	.37	.12
Be/Do (Do)	Mean	.71	.24
	SD	.34	.33
	r	.35	.19

Relationship Between the Rice/Wexler Test of Early Grammatical Impairment and Word Classes Subtest From *CELF®–3*

Table 5.7 reports the means, standard deviations, and correlations between *CELF®–3* Word Classes and Rice/Wexler probes and Elicited Grammar Composite for children ages 6.00–8.11.

The range for the children in the normal language group is .13 to .25, which is lower than the parallel association with Word Structure subtests which ranged from .37 to .53. For children with language impairments, the range is .06 to .14, which is lower than the parallel association with Word Structure ranging from .41 to .54.

Table 5.7 Means, Standard Deviations, and Correlations Between *CELF®–3* Word Classes and Rice/Wexler Probes and Elicited Grammar Composite for Children with Language Disorders

Subtest		Normal Language Group Ages 6.00–6.11 (n = 149)	Language Disorder Group Ages 6.00–8.11 (n = 279)
Word Classes	Mean	11.23	7.90
	SD	2.68	2.62
Elicited Grammar Composite	Mean	.91	.67
	SD	.12	.25
	r	.25	.13
Third Person Singular	Mean	.92	.68
	SD	.18	.35
	r	.22	.09
Past Tense	Mean	.91	.69
	SD	.14	.26
	r	.13	.06
Be/Do (Be)	Mean	.94	.72
	SD	.10	.26
	r	.21	.14
Be/Do (Do)	Mean	.87	.57
	SD	.16	.33
	r	.23	.12

As expected, overall the correlations between the Basic Concepts and Word Classes subtests and the Rice/Wexler are low, indicating that both of these subtests measure a different construct than the Rice/Wexler. This pattern remains consistent for both the language impairment group and the normal language group. At the same time, the lower associations between the Elicited Grammar Composite and Word Classes/Basic Concepts for the children with language impairments versus the children in the normal language group may reflect a tighter coherence within the linguistic system for the children in the normal group than what exists for the children with language impairments. It is consistent with the possibility that the finiteness markers measured in Rice/Wexler tend to fall behind other elements of language acquisition for the affected children.

Relationship Between the *Rice/Wexler Test of Early Grammatical Impairment* and Non-Verbal IQ Scores

Another study was conducted to investigate the relationship between Rice/Wexler scores and non-verbal IQ scores, and the ability to identify children who may be identified as SLI on the basis of performance on the Rice/Wexler.

All children included in this study were part of the language disordered study from the standardization research. To be included in this study, the children had to have a non-verbal IQ score of 85 or above and have no other known handicapping conditions or diagnoses, in addition to having a diagnosed language disorder. The IQ information was obtained from the child's academic records if IQ testing had been completed within the past 12 months. In some cases, an IQ test (with a non-verbal component) was administered and the score was obtained specifically for this study. Both parent reports and clinician reports were used to determine whether or not the child had any other handicapping conditions. According to accepted definitions, these children are considered to meet the criteria for Specific Language Impairment.

A total of sixty-nine children between the ages of 3.00–8.11 were included in this study. There were 68% males and 32% females; the racial/ethnic representation was 65% White, 20% African American, 9% Hispanic, and 6% other races.

Table 5.8 reports the means, standard deviations, and correlations for the Elicited Grammar Composite and the non-verbal IQ scores.

Table 5.8 Means, Standard Deviations, and Correlations for the Elicited Grammar Composite and Non-Verbal IQ Scores for Children with Language Disorders

	Mean	SD	Correlation
Non-Verbal IQ Score	101.75	12.09	-.17
Elicited Grammar Composite	.63	.27	

As expected, for children with language disorders there is a very low correlation between the non-verbal IQ score and performance on the Rice/Wexler. As shown by the earlier research studies, performance on these finiteness measures is not predicted by a child's nonverbal IQ level. This would be the expected pattern for children who are considered Specific Language Impaired; that is, children whose nonverbal IQ is in the normal range, and have no other impairments but show language deficits. Although further research on the effects of Specific Language Impairment on children is needed and encouraged, as is research on language impairments in children with borderline normal or below normal nonverbal IQ levels, initial findings from this study indicate that the Rice/Wexler may be a useful tool in identifying children with language impairments across different levels of nonverbal IQ.

Relationship Between Rice/Wexler and Parent Education Levels

Table 5.9 reports the means, standard deviations, and correlations for the Elicited Grammar Composite and the Grammaticality Judgment A' scores for children with language disorders (n = 444) and children in the normal language group (n = 393), collapsed across age levels. As expected from previous studies, the correlations are very low, suggesting that parent education level does not predict children's performance on the grammatical marker. This is an important outcome because it shows that children's performance on the marker is not confounded with parent education level. The Rice/Wexler may be a useful tool in identifying children with language impairments across different parent education levels.

Table 5.9 Means, Standard Deviations, and Correlations for the Elicited Grammar Composite and Grammaticality Judgments and Parent Education Levels

		Normal Language Group (n = 393)	Language Disorder Group (n = 444)
Parent Education	M	2.98	2.58
	SD	1.02	.97
Elicited Grammar Composite	M	.86	.50
	SD	.16	.28
	r	.03	.06
Dropped Marker A'	M	.63	.54
	SD	.39	.31
	r	-.01	-.08
Agreement A'	M	.66	.60
	SD	.40	.34
	r	-.02	-.06
Dropped -ing A'	M	.67	.61
	SD	.41	.37
	r	-.04	-.08

Descriptive Data

Table 4.13 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Normal Language Group

		Third Person Singular Probe		Past Tense Probe		Be/Do Probe (Be Score)		Be/Do Probe (Do Score)		Elicited Grammar Composite		Dropped Marker*		Agreement*		Dropped -ing*	
Age	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3.00–3.05	43	.71	.30	.65	.25	.72	.28	.60	.35	.67	.22						
3.06–3.11	50	.80	.27	.78	.22	.86	.17	.71	.36	.79	.17						
4.00–4.05	50	.87	.24	.84	.19	.87	.18	.74	.34	.83	.20	.70	.25	.75	.25	.73	.31
4.06–4.11	50	.91	.18	.90	.10	.90	.16	.83	.23	.89	.11	.75	.21	.81	.20	.83	.22
5.00–5.05	50	.93	.11	.88	.12	.93	.12	.87	.18	.90	.10	.80	.20	.84	.21	.85	.24
5.06–5.11	50	.97	.06	.93	.08	.93	.08	.83	.20	.92	.08	.83	.18	.87	.16	.87	.17
6.00–6.05	50	.97	.07	.93	.08	.96	.06	.90	.13	.94	.06	.92	.12	.94	.12	.94	.15
6.06–6.11	50	.96	.07	.94	.06	.96	.06	.90	.14	.94	.06	.93	.09	.98	.05	.97	.07

* As described in text, these scores represent A' calculations. All other scores represent a percentage correct.

Table 4.14 Means and Standard Deviations for the Rice/Wexler Probe Scores, Elicited Grammar Composite, and Grammaticality Judgment Scores by Age—Children in the Language Disorder Group

		Third Person Singular Probe		Past Tense Probe		Be/Do Probe (Be Score)		Be/Do Probe (Do Score)		Elicited Grammar Composite		Dropped Marker*		Agreement*		Dropped -ing*	
Age	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3.00–3.05	20	.29	.34	.36	.31	.23	.37	.14	.33	.25	.20						
3.06–3.11	24	.26	.35	.30	.25	.40	.35	.09	.24	.26	.21						
4.00–4.05	50	.38	.34	.38	.30	.48	.32	.21	.32	.36	.24	.43	.30	.46	.35	.45	.38
4.06–4.11	50	.39	.34	.48	.27	.57	.34	.20	.31	.41	.23	.43	.29	.50	.34	.53	.38
5.00–5.05	50	.47	.36	.44	.31	.46	.31	.25	.31	.41	.26	.53	.22	.60	.25	.64	.30
5.06–5.11	50	.47	.35	.49	.30	.60	.25	.30	.28	.47	.24	.58	.20	.65	.25	.69	.29
6.00–6.05	50	.57	.37	.60	.24	.59	.28	.36	.35	.53	.24	.58	.26	.65	.29	.61	.34
6.06–6.11	50	.57	.34	.58	.28	.62	.28	.44	.32	.55	.25	.63	.23	.73	.20	.71	.27
7.00–7.11	50	.69	.34	.76	.23	.79	.23	.67	.31	.73	.21	.76	.20	.82	.20	.84	.24
8.00–8.11	50	.73	.35	.78	.24	.78	.22	.67	.27	.74	.22	.83	.16	.88	.15	.92	.14

* As described in text, these scores represent A' calculations. All other scores represent a percentage correct.

Criterion Scores

A

Third Person Singular Probe

Age	Criterion Score
3.00–3.05	51
3.06–3.11	51
4.00–4.05	76
4.06–4.11	76
5.00–5.05	89
5.06–5.11	81
6.00–6.05	92
6.06–6.11	91
7.00–7.11	*90
8.00–8.11	**90

* 68% of children in the language disorder group scored at or below this level

** 54% of children in the language disorder group scored at or below this level

Past Tense Probe

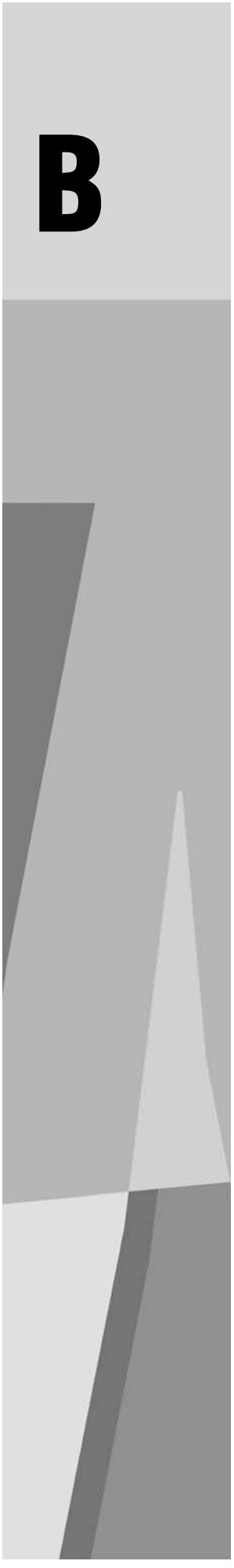
Age	Criterion Score
3.00–3.05	68
3.06–3.11	60
4.00–4.05	68
4.06–4.11	73
5.00–5.05	73
5.06–5.11	79
6.00–6.05	81
6.06–6.11	87
7.00–7.11	94
8.00–8.11	94

Screener

Age	Criterion Score
3.00–3.05	47
3.06–3.11	62
4.00–4.05	63
4.06–4.11	65
5.00–5.05	78
5.06–5.11	80
6.00–6.05	85
6.06–6.11	88
7.00–7.11	94
8.00–8.11	97

Sensitivity and Specificity

B



Ages 3.00–3.05

Score	Specificity n = 43	Sensitivity n = 20	Score	Specificity n = 43	Sensitivity n = 20
0.00	1.00	0.00	0.51	0.74	0.80
0.01	0.93	0.40	0.52	0.74	0.80
0.02	0.93	0.40	0.53	0.74	0.80
0.03	0.93	0.40	0.54	0.74	0.80
0.04	0.93	0.40	0.55	0.74	0.80
0.05	0.93	0.40	0.56	0.74	0.80
0.06	0.93	0.40	0.57	0.72	0.80
0.07	0.93	0.40	0.58	0.70	0.80
0.08	0.93	0.40	0.59	0.70	0.80
0.09	0.93	0.40	0.60	0.70	0.80
0.10	0.93	0.40	0.61	0.67	0.80
0.11	0.93	0.40	0.62	0.67	0.80
0.12	0.93	0.40	0.63	0.67	0.80
0.13	0.93	0.40	0.64	0.65	0.85
0.14	0.91	0.40	0.65	0.65	0.85
0.15	0.91	0.40	0.66	0.65	0.85
0.16	0.91	0.40	0.67	0.65	0.85
0.17	0.91	0.40	0.68	0.63	0.85
0.18	0.91	0.45	0.69	0.63	0.85
0.19	0.91	0.45	0.70	0.63	0.85
0.20	0.91	0.45	0.71	0.58	0.85
0.21	0.91	0.55	0.72	0.56	0.85
0.22	0.91	0.55	0.73	0.56	0.85
0.23	0.91	0.55	0.74	0.56	0.85
0.24	0.91	0.55	0.75	0.56	0.85
0.25	0.91	0.55	0.76	0.56	0.85
0.26	0.88	0.60	0.77	0.56	0.85
0.27	0.88	0.60	0.78	0.56	0.85
0.28	0.88	0.60	0.79	0.53	0.85
0.29	0.88	0.60	0.80	0.53	0.85
0.30	0.88	0.60	0.81	0.49	0.90
0.31	0.88	0.60	0.82	0.49	0.90
0.32	0.88	0.60	0.83	0.49	0.90
0.33	0.88	0.60	0.84	0.49	0.90
0.34	0.88	0.70	0.85	0.49	0.90
0.35	0.88	0.70	0.86	0.49	0.90
0.36	0.88	0.70	0.87	0.49	0.90
0.37	0.88	0.70	0.88	0.49	0.90
0.38	0.88	0.70	0.89	0.42	0.90
0.39	0.86	0.70	0.90	0.42	0.90
0.40	0.86	0.70	0.91	0.28	0.90
0.41	0.86	0.75	0.92	0.28	0.90
0.42	0.86	0.75	0.93	0.28	0.90
0.43	0.86	0.75	0.94	0.28	0.90
0.44	0.81	0.75	0.95	0.28	0.90
0.45	0.81	0.75	0.96	0.28	0.90
0.46	0.81	0.75	0.97	0.28	0.90
0.47	0.81	0.75	0.98	0.28	0.90
0.48	0.81	0.75	0.99	0.28	0.90
0.49	0.81	0.75	1.00	0.28	0.90
0.50	0.81	0.75			

Ages 3.06–3.11

Score	Specificity n = 50	Sensitivity n = 24	Score	Specificity n = 50	Sensitivity n = 24
0.00	1.00	0.00	0.51	0.80	0.83
0.01	0.98	0.38	0.52	0.80	0.83
0.02	0.98	0.38	0.53	0.80	0.83
0.03	0.98	0.38	0.54	0.80	0.83
0.04	0.98	0.38	0.55	0.80	0.83
0.05	0.98	0.38	0.56	0.80	0.83
0.06	0.98	0.38	0.57	0.80	0.83
0.07	0.98	0.38	0.58	0.80	0.83
0.08	0.98	0.38	0.59	0.80	0.83
0.09	0.98	0.38	0.60	0.80	0.83
0.10	0.98	0.38	0.61	0.78	0.83
0.11	0.98	0.38	0.62	0.78	0.83
0.12	0.98	0.46	0.63	0.78	0.83
0.13	0.98	0.46	0.64	0.76	0.83
0.14	0.96	0.46	0.65	0.76	0.83
0.15	0.94	0.46	0.66	0.76	0.83
0.16	0.94	0.46	0.67	0.76	0.83
0.17	0.94	0.46	0.68	0.70	0.83
0.18	0.94	0.58	0.69	0.70	0.83
0.19	0.94	0.58	0.70	0.70	0.83
0.20	0.94	0.58	0.71	0.70	0.83
0.21	0.94	0.67	0.72	0.70	0.83
0.22	0.94	0.67	0.73	0.70	0.83
0.23	0.94	0.75	0.74	0.70	0.83
0.24	0.94	0.75	0.75	0.70	0.83
0.25	0.94	0.75	0.76	0.70	0.83
0.26	0.92	0.75	0.77	0.70	0.83
0.27	0.92	0.75	0.78	0.70	0.83
0.28	0.92	0.75	0.79	0.68	0.83
0.29	0.92	0.75	0.80	0.68	0.83
0.30	0.92	0.75	0.81	0.64	0.83
0.31	0.92	0.79	0.82	0.64	0.83
0.32	0.92	0.79	0.83	0.64	0.83
0.33	0.92	0.79	0.84	0.62	0.83
0.34	0.92	0.79	0.85	0.62	0.83
0.35	0.92	0.79	0.86	0.62	0.83
0.36	0.92	0.79	0.87	0.60	0.83
0.37	0.92	0.79	0.88	0.60	0.83
0.38	0.92	0.79	0.89	0.56	0.83
0.39	0.92	0.79	0.90	0.52	0.88
0.40	0.92	0.79	0.91	0.48	0.88
0.41	0.90	0.79	0.92	0.48	0.88
0.42	0.90	0.79	0.93	0.48	0.88
0.43	0.90	0.79	0.94	0.48	0.88
0.44	0.90	0.79	0.95	0.48	0.88
0.45	0.88	0.79	0.96	0.48	0.88
0.46	0.88	0.79	0.97	0.48	0.88
0.47	0.88	0.79	0.98	0.48	0.88
0.48	0.88	0.79	0.99	0.48	0.88
0.49	0.88	0.79	1.00	0.48	0.88
0.50	0.88	0.79			

Ages 4.00–4.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	0.92	0.66
0.01	0.96	0.26	0.52	0.92	0.66
0.02	0.96	0.26	0.53	0.92	0.66
0.03	0.96	0.26	0.54	0.92	0.66
0.04	0.96	0.26	0.55	0.92	0.66
0.05	0.96	0.26	0.56	0.92	0.66
0.06	0.96	0.26	0.57	0.90	0.68
0.07	0.96	0.26	0.58	0.90	0.68
0.08	0.96	0.26	0.59	0.90	0.68
0.09	0.96	0.26	0.60	0.90	0.68
0.10	0.96	0.26	0.61	0.90	0.70
0.11	0.96	0.26	0.62	0.90	0.70
0.12	0.96	0.26	0.63	0.90	0.70
0.13	0.96	0.26	0.64	0.90	0.70
0.14	0.96	0.28	0.65	0.90	0.70
0.15	0.96	0.32	0.66	0.90	0.70
0.16	0.96	0.32	0.67	0.90	0.70
0.17	0.96	0.32	0.68	0.84	0.74
0.18	0.96	0.34	0.69	0.84	0.74
0.19	0.96	0.34	0.70	0.84	0.74
0.20	0.96	0.34	0.71	0.82	0.76
0.21	0.96	0.38	0.72	0.82	0.76
0.22	0.96	0.38	0.73	0.82	0.76
0.23	0.94	0.42	0.74	0.82	0.76
0.24	0.94	0.42	0.75	0.82	0.76
0.25	0.94	0.42	0.76	0.78	0.80
0.26	0.94	0.50	0.77	0.78	0.80
0.27	0.94	0.50	0.78	0.78	0.80
0.28	0.94	0.50	0.79	0.76	0.80
0.29	0.94	0.50	0.80	0.76	0.80
0.30	0.94	0.50	0.81	0.74	0.86
0.31	0.94	0.52	0.82	0.74	0.86
0.32	0.94	0.52	0.83	0.74	0.86
0.33	0.94	0.52	0.84	0.74	0.86
0.34	0.94	0.58	0.85	0.74	0.86
0.35	0.94	0.58	0.86	0.74	0.86
0.36	0.94	0.58	0.87	0.72	0.86
0.37	0.94	0.58	0.88	0.72	0.86
0.38	0.94	0.58	0.89	0.68	0.88
0.39	0.94	0.60	0.90	0.62	0.88
0.40	0.94	0.60	0.91	0.62	0.92
0.41	0.94	0.64	0.92	0.62	0.92
0.42	0.94	0.64	0.93	0.62	0.92
0.43	0.94	0.64	0.94	0.62	0.92
0.44	0.94	0.64	0.95	0.62	0.92
0.45	0.94	0.66	0.96	0.62	0.92
0.46	0.94	0.66	0.97	0.62	0.92
0.47	0.94	0.66	0.98	0.62	0.92
0.48	0.94	0.66	0.99	0.62	0.92
0.49	0.94	0.66	1.00	0.62	0.92
0.50	0.94	0.66			

Ages 4.06–4.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	0.96	0.68
0.01	0.98	0.26	0.52	0.96	0.68
0.02	0.98	0.26	0.53	0.96	0.68
0.03	0.98	0.26	0.54	0.96	0.68
0.04	0.98	0.26	0.55	0.96	0.68
0.05	0.98	0.26	0.56	0.96	0.68
0.06	0.98	0.26	0.57	0.94	0.70
0.07	0.98	0.26	0.58	0.94	0.72
0.08	0.98	0.26	0.59	0.94	0.72
0.09	0.98	0.26	0.60	0.94	0.72
0.10	0.98	0.26	0.61	0.94	0.72
0.11	0.98	0.30	0.62	0.94	0.72
0.12	0.98	0.30	0.63	0.94	0.72
0.13	0.98	0.30	0.64	0.92	0.76
0.14	0.98	0.32	0.65	0.92	0.76
0.15	0.98	0.32	0.66	0.92	0.76
0.16	0.98	0.32	0.67	0.92	0.76
0.17	0.98	0.32	0.68	0.92	0.76
0.18	0.98	0.34	0.69	0.92	0.76
0.19	0.98	0.34	0.70	0.92	0.76
0.20	0.98	0.34	0.71	0.92	0.76
0.21	0.98	0.40	0.72	0.92	0.78
0.22	0.98	0.40	0.73	0.92	0.78
0.23	0.98	0.40	0.74	0.92	0.78
0.24	0.98	0.40	0.75	0.92	0.78
0.25	0.98	0.40	0.76	0.90	0.80
0.26	0.98	0.42	0.77	0.90	0.80
0.27	0.98	0.42	0.78	0.90	0.80
0.28	0.98	0.42	0.79	0.88	0.80
0.29	0.98	0.42	0.80	0.88	0.80
0.30	0.98	0.42	0.81	0.80	0.86
0.31	0.98	0.42	0.82	0.80	0.86
0.32	0.98	0.42	0.83	0.80	0.86
0.33	0.98	0.42	0.84	0.80	0.86
0.34	0.98	0.46	0.85	0.80	0.86
0.35	0.98	0.46	0.86	0.80	0.86
0.36	0.98	0.46	0.87	0.80	0.86
0.37	0.98	0.46	0.88	0.80	0.86
0.38	0.98	0.46	0.89	0.80	0.88
0.39	0.98	0.54	0.90	0.74	0.88
0.40	0.98	0.54	0.91	0.60	0.94
0.41	0.98	0.54	0.92	0.60	0.94
0.42	0.98	0.54	0.93	0.60	0.94
0.43	0.98	0.54	0.94	0.60	0.94
0.44	0.98	0.58	0.95	0.60	0.94
0.45	0.98	0.60	0.96	0.60	0.94
0.46	0.98	0.60	0.97	0.60	0.94
0.47	0.98	0.60	0.98	0.60	0.94
0.48	0.98	0.60	0.99	0.60	0.94
0.49	0.98	0.60	1.00	0.60	0.94
0.50	0.98	0.60			

Ages 5.00–5.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.62
0.01	1.00	0.14	0.52	1.00	0.62
0.02	1.00	0.14	0.53	1.00	0.62
0.03	1.00	0.14	0.54	1.00	0.62
0.04	1.00	0.14	0.55	1.00	0.62
0.05	1.00	0.14	0.56	1.00	0.62
0.06	1.00	0.14	0.57	0.98	0.64
0.07	1.00	0.14	0.58	0.98	0.64
0.08	1.00	0.14	0.59	0.98	0.64
0.09	1.00	0.14	0.60	0.98	0.64
0.10	1.00	0.14	0.61	0.98	0.64
0.11	1.00	0.18	0.62	0.98	0.64
0.12	1.00	0.22	0.63	0.98	0.64
0.13	1.00	0.22	0.64	0.98	0.64
0.14	1.00	0.22	0.65	0.98	0.64
0.15	1.00	0.26	0.66	0.98	0.64
0.16	1.00	0.26	0.67	0.98	0.64
0.17	1.00	0.26	0.68	0.94	0.64
0.18	1.00	0.30	0.69	0.94	0.64
0.19	1.00	0.30	0.70	0.94	0.64
0.20	1.00	0.30	0.71	0.88	0.64
0.21	1.00	0.32	0.72	0.88	0.64
0.22	1.00	0.32	0.73	0.88	0.64
0.23	1.00	0.38	0.74	0.88	0.64
0.24	1.00	0.38	0.75	0.88	0.64
0.25	1.00	0.38	0.76	0.88	0.66
0.26	1.00	0.40	0.77	0.88	0.66
0.27	1.00	0.40	0.78	0.88	0.66
0.28	1.00	0.40	0.79	0.88	0.72
0.29	1.00	0.40	0.80	0.88	0.72
0.30	1.00	0.40	0.81	0.84	0.76
0.31	1.00	0.42	0.82	0.84	0.76
0.32	1.00	0.42	0.83	0.84	0.76
0.33	1.00	0.42	0.84	0.82	0.76
0.34	1.00	0.50	0.85	0.82	0.76
0.35	1.00	0.50	0.86	0.82	0.76
0.36	1.00	0.50	0.87	0.82	0.78
0.37	1.00	0.50	0.88	0.82	0.78
0.38	1.00	0.50	0.89	0.78	0.80
0.39	1.00	0.50	0.90	0.74	0.82
0.40	1.00	0.50	0.91	0.66	0.84
0.41	1.00	0.52	0.92	0.66	0.84
0.42	1.00	0.52	0.93	0.66	0.84
0.43	1.00	0.52	0.94	0.66	0.84
0.44	1.00	0.54	0.95	0.66	0.84
0.45	1.00	0.54	0.96	0.66	0.84
0.46	1.00	0.54	0.97	0.66	0.84
0.47	1.00	0.54	0.98	0.66	0.84
0.48	1.00	0.54	0.99	0.66	0.84
0.49	1.00	0.54	1.00	0.66	0.84
0.50	1.00	0.54			

Ages 5.06–5.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.62
0.01	1.00	0.16	0.52	1.00	0.62
0.02	1.00	0.16	0.53	1.00	0.62
0.03	1.00	0.16	0.54	1.00	0.62
0.04	1.00	0.16	0.55	1.00	0.62
0.05	1.00	0.16	0.56	1.00	0.62
0.06	1.00	0.16	0.57	1.00	0.64
0.07	1.00	0.16	0.58	1.00	0.66
0.08	1.00	0.16	0.59	1.00	0.66
0.09	1.00	0.16	0.60	1.00	0.66
0.10	1.00	0.16	0.61	1.00	0.66
0.11	1.00	0.22	0.62	1.00	0.66
0.12	1.00	0.22	0.63	1.00	0.66
0.13	1.00	0.22	0.64	1.00	0.66
0.14	1.00	0.22	0.65	1.00	0.66
0.15	1.00	0.24	0.66	1.00	0.66
0.16	1.00	0.24	0.67	1.00	0.66
0.17	1.00	0.24	0.68	1.00	0.68
0.18	1.00	0.26	0.69	1.00	0.68
0.19	1.00	0.26	0.70	1.00	0.68
0.20	1.00	0.26	0.71	0.98	0.70
0.21	1.00	0.28	0.72	0.98	0.70
0.22	1.00	0.28	0.73	0.98	0.70
0.23	1.00	0.36	0.74	0.98	0.70
0.24	1.00	0.36	0.75	0.98	0.70
0.25	1.00	0.36	0.76	0.98	0.72
0.26	1.00	0.36	0.77	0.98	0.72
0.27	1.00	0.36	0.78	0.98	0.72
0.28	1.00	0.36	0.79	0.98	0.74
0.29	1.00	0.36	0.80	0.98	0.74
0.30	1.00	0.38	0.81	0.96	0.80
0.31	1.00	0.38	0.82	0.96	0.80
0.32	1.00	0.38	0.83	0.96	0.80
0.33	1.00	0.38	0.84	0.96	0.82
0.34	1.00	0.40	0.85	0.96	0.82
0.35	1.00	0.40	0.86	0.96	0.82
0.36	1.00	0.40	0.87	0.96	0.82
0.37	1.00	0.40	0.88	0.96	0.82
0.38	1.00	0.40	0.89	0.94	0.82
0.39	1.00	0.40	0.90	0.90	0.82
0.40	1.00	0.40	0.91	0.76	0.86
0.41	1.00	0.50	0.92	0.76	0.86
0.42	1.00	0.50	0.93	0.76	0.86
0.43	1.00	0.50	0.94	0.76	0.86
0.44	1.00	0.50	0.95	0.76	0.86
0.45	1.00	0.52	0.96	0.76	0.86
0.46	1.00	0.52	0.97	0.76	0.86
0.47	1.00	0.52	0.98	0.76	0.86
0.48	1.00	0.52	0.99	0.76	0.86
0.49	1.00	0.52	1.00	0.76	0.86
0.50	1.00	0.52			

Ages 6.00–6.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.38
0.01	1.00	0.18	0.52	1.00	0.38
0.02	1.00	0.18	0.53	1.00	0.38
0.03	1.00	0.18	0.54	1.00	0.38
0.04	1.00	0.18	0.55	1.00	0.38
0.05	1.00	0.18	0.56	1.00	0.38
0.06	1.00	0.18	0.57	1.00	0.42
0.07	1.00	0.18	0.58	1.00	0.42
0.08	1.00	0.18	0.59	1.00	0.42
0.09	1.00	0.18	0.60	1.00	0.42
0.10	1.00	0.18	0.61	0.98	0.46
0.11	1.00	0.24	0.62	0.98	0.46
0.12	1.00	0.24	0.63	0.98	0.46
0.13	1.00	0.24	0.64	0.98	0.50
0.14	1.00	0.24	0.65	0.98	0.50
0.15	1.00	0.24	0.66	0.98	0.50
0.16	1.00	0.24	0.67	0.98	0.50
0.17	1.00	0.24	0.68	0.98	0.54
0.18	1.00	0.24	0.69	0.98	0.54
0.19	1.00	0.24	0.70	0.98	0.54
0.20	1.00	0.24	0.71	0.98	0.56
0.21	1.00	0.26	0.72	0.98	0.58
0.22	1.00	0.26	0.73	0.98	0.58
0.23	1.00	0.26	0.74	0.98	0.58
0.24	1.00	0.26	0.75	0.98	0.58
0.25	1.00	0.26	0.76	0.98	0.62
0.26	1.00	0.26	0.77	0.98	0.62
0.27	1.00	0.26	0.78	0.98	0.62
0.28	1.00	0.26	0.79	0.98	0.66
0.29	1.00	0.26	0.80	0.98	0.66
0.30	1.00	0.26	0.81	0.94	0.70
0.31	1.00	0.28	0.82	0.94	0.70
0.32	1.00	0.28	0.83	0.94	0.70
0.33	1.00	0.28	0.84	0.94	0.72
0.34	1.00	0.30	0.85	0.94	0.72
0.35	1.00	0.30	0.86	0.94	0.72
0.36	1.00	0.30	0.87	0.94	0.72
0.37	1.00	0.30	0.88	0.94	0.72
0.38	1.00	0.30	0.89	0.94	0.72
0.39	1.00	0.32	0.90	0.88	0.74
0.40	1.00	0.32	0.91	0.82	0.80
0.41	1.00	0.32	0.92	0.82	0.80
0.42	1.00	0.32	0.93	0.82	0.80
0.43	1.00	0.32	0.94	0.82	0.80
0.44	1.00	0.32	0.95	0.82	0.80
0.45	1.00	0.36	0.96	0.82	0.80
0.46	1.00	0.36	0.97	0.82	0.80
0.47	1.00	0.36	0.98	0.82	0.80
0.48	1.00	0.36	0.99	0.82	0.80
0.49	1.00	0.36	1.00	0.82	0.80
0.50	1.00	0.36			

Ages 6.06–6.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.48
0.01	1.00	0.14	0.52	1.00	0.48
0.02	1.00	0.14	0.53	1.00	0.48
0.03	1.00	0.14	0.54	1.00	0.48
0.04	1.00	0.14	0.55	1.00	0.48
0.05	1.00	0.14	0.56	1.00	0.48
0.06	1.00	0.14	0.57	1.00	0.50
0.07	1.00	0.14	0.58	1.00	0.50
0.08	1.00	0.14	0.59	1.00	0.50
0.09	1.00	0.14	0.60	1.00	0.50
0.10	1.00	0.14	0.61	1.00	0.50
0.11	1.00	0.16	0.62	1.00	0.50
0.12	1.00	0.18	0.63	1.00	0.50
0.13	1.00	0.18	0.64	1.00	0.50
0.14	1.00	0.18	0.65	1.00	0.50
0.15	1.00	0.18	0.66	1.00	0.50
0.16	1.00	0.18	0.67	1.00	0.50
0.17	1.00	0.18	0.68	1.00	0.54
0.18	1.00	0.18	0.69	1.00	0.54
0.19	1.00	0.18	0.70	1.00	0.54
0.20	1.00	0.18	0.71	1.00	0.56
0.21	1.00	0.20	0.72	1.00	0.56
0.22	1.00	0.20	0.73	1.00	0.56
0.23	1.00	0.20	0.74	1.00	0.56
0.24	1.00	0.20	0.75	1.00	0.56
0.25	1.00	0.20	0.76	0.96	0.58
0.26	1.00	0.22	0.77	0.96	0.58
0.27	1.00	0.22	0.78	0.96	0.58
0.28	1.00	0.22	0.79	0.94	0.60
0.29	1.00	0.22	0.80	0.94	0.60
0.30	1.00	0.24	0.81	0.92	0.72
0.31	1.00	0.24	0.82	0.92	0.72
0.32	1.00	0.24	0.83	0.92	0.72
0.33	1.00	0.24	0.84	0.92	0.74
0.34	1.00	0.26	0.85	0.92	0.74
0.35	1.00	0.26	0.86	0.92	0.74
0.36	1.00	0.26	0.87	0.92	0.74
0.37	1.00	0.26	0.88	0.92	0.74
0.38	1.00	0.26	0.89	0.92	0.76
0.39	1.00	0.26	0.90	0.84	0.78
0.40	1.00	0.26	0.91	0.68	0.84
0.41	1.00	0.34	0.92	0.68	0.84
0.42	1.00	0.34	0.93	0.68	0.84
0.43	1.00	0.34	0.94	0.68	0.84
0.44	1.00	0.34	0.95	0.68	0.84
0.45	1.00	0.38	0.96	0.68	0.84
0.46	1.00	0.38	0.97	0.68	0.84
0.47	1.00	0.38	0.98	0.68	0.84
0.48	1.00	0.38	0.99	0.68	0.84
0.49	1.00	0.38	1.00	0.68	0.84
0.50	1.00	0.38			

Ages 7.00–7.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.10	0.51	0.24
0.01	0.10	0.52	0.24
0.02	0.10	0.53	0.24
0.03	0.10	0.54	0.24
0.04	0.10	0.55	0.24
0.05	0.10	0.56	0.26
0.06	0.10	0.57	0.26
0.07	0.10	0.58	0.26
0.08	0.10	0.59	0.26
0.09	0.10	0.60	0.32
0.10	0.12	0.61	0.32
0.11	0.12	0.62	0.32
0.12	0.12	0.63	0.34
0.13	0.12	0.64	0.34
0.14	0.12	0.65	0.34
0.15	0.12	0.66	0.34
0.16	0.12	0.67	0.38
0.17	0.12	0.68	0.38
0.18	0.12	0.69	0.38
0.19	0.12	0.70	0.44
0.20	0.18	0.71	0.44
0.21	0.18	0.72	0.44
0.22	0.20	0.73	0.44
0.23	0.20	0.74	0.44
0.24	0.20	0.75	0.44
0.25	0.20	0.76	0.44
0.26	0.20	0.77	0.44
0.27	0.20	0.78	0.46
0.28	0.20	0.79	0.46
0.29	0.20	0.80	0.54
0.30	0.20	0.81	0.54
0.31	0.20	0.82	0.54
0.32	0.20	0.83	0.54
0.33	0.22	0.84	0.54
0.34	0.22	0.85	0.54
0.35	0.22	0.86	0.54
0.36	0.22	0.87	0.54
0.37	0.22	0.88	0.56
0.38	0.22	0.89	0.58
0.39	0.22	0.90	0.68
0.40	0.22	0.91	0.68
0.41	0.22	0.92	0.68
0.42	0.22	0.93	0.68
0.43	0.22	0.94	0.68
0.44	0.22	0.95	0.68
0.45	0.22	0.96	0.68
0.46	0.22	0.97	0.68
0.47	0.22	0.98	0.68
0.48	0.22	0.99	0.68
0.49	0.22	1.00	1.00
0.50	0.24		

Ages 8.00–8.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.08	0.51	0.24
0.01	0.08	0.52	0.24
0.02	0.08	0.53	0.24
0.03	0.08	0.54	0.24
0.04	0.08	0.55	0.24
0.05	0.08	0.56	0.24
0.06	0.08	0.57	0.24
0.07	0.08	0.58	0.24
0.08	0.08	0.59	0.24
0.09	0.08	0.60	0.28
0.10	0.10	0.61	0.28
0.11	0.10	0.62	0.28
0.12	0.10	0.63	0.28
0.13	0.14	0.64	0.28
0.14	0.14	0.65	0.28
0.15	0.14	0.66	0.28
0.16	0.14	0.67	0.34
0.17	0.14	0.68	0.34
0.18	0.14	0.69	0.34
0.19	0.14	0.70	0.38
0.20	0.16	0.71	0.38
0.21	0.16	0.72	0.38
0.22	0.18	0.73	0.38
0.23	0.18	0.74	0.38
0.24	0.18	0.75	0.38
0.25	0.18	0.76	0.38
0.26	0.18	0.77	0.38
0.27	0.18	0.78	0.38
0.28	0.18	0.79	0.38
0.29	0.18	0.80	0.42
0.30	0.20	0.81	0.42
0.31	0.20	0.82	0.42
0.32	0.20	0.83	0.42
0.33	0.20	0.84	0.42
0.34	0.20	0.85	0.42
0.35	0.20	0.86	0.42
0.36	0.20	0.87	0.42
0.37	0.20	0.88	0.44
0.38	0.20	0.89	0.48
0.39	0.20	0.90	0.54
0.40	0.24	0.91	0.54
0.41	0.24	0.92	0.54
0.42	0.24	0.93	0.54
0.43	0.24	0.94	0.54
0.44	0.24	0.95	0.54
0.45	0.24	0.96	0.54
0.46	0.24	0.97	0.54
0.47	0.24	0.98	0.54
0.48	0.24	0.99	0.54
0.49	0.24	1.00	1.00
0.50	0.24		

Ages 3.00–3.05

Score	Specificity n = 43	Sensitivity n = 20	Score	Specificity n = 43	Sensitivity n = 20
0.00	1.00	0.00	0.51	0.70	0.70
0.01	1.00	0.20	0.52	0.70	0.70
0.02	1.00	0.20	0.53	0.70	0.70
0.03	1.00	0.20	0.54	0.67	0.70
0.04	1.00	0.20	0.55	0.67	0.70
0.05	1.00	0.20	0.56	0.67	0.70
0.06	1.00	0.20	0.57	0.67	0.70
0.07	1.00	0.20	0.58	0.67	0.70
0.08	1.00	0.20	0.59	0.65	0.70
0.09	1.00	0.20	0.60	0.65	0.70
0.10	1.00	0.20	0.61	0.65	0.75
0.11	1.00	0.20	0.62	0.58	0.75
0.12	1.00	0.20	0.63	0.58	0.75
0.13	1.00	0.20	0.64	0.56	0.75
0.14	0.98	0.25	0.65	0.56	0.75
0.15	0.98	0.25	0.66	0.56	0.75
0.16	0.98	0.25	0.67	0.56	0.75
0.17	0.98	0.25	0.68	0.53	0.80
0.18	0.93	0.30	0.69	0.53	0.80
0.19	0.93	0.30	0.70	0.51	0.80
0.20	0.93	0.30	0.71	0.51	0.80
0.21	0.91	0.35	0.72	0.49	0.85
0.22	0.91	0.35	0.73	0.44	0.85
0.23	0.91	0.45	0.74	0.40	0.90
0.24	0.91	0.45	0.75	0.40	0.90
0.25	0.91	0.45	0.76	0.40	0.90
0.26	0.91	0.50	0.77	0.37	0.90
0.27	0.91	0.50	0.78	0.35	0.90
0.28	0.91	0.50	0.79	0.30	0.90
0.29	0.88	0.50	0.80	0.28	0.90
0.30	0.86	0.55	0.81	0.26	0.90
0.31	0.86	0.55	0.82	0.26	0.90
0.32	0.86	0.55	0.83	0.26	0.90
0.33	0.86	0.55	0.84	0.23	0.90
0.34	0.86	0.60	0.85	0.23	0.90
0.35	0.86	0.60	0.86	0.23	0.90
0.36	0.86	0.60	0.87	0.23	0.90
0.37	0.84	0.65	0.88	0.23	0.90
0.38	0.84	0.65	0.89	0.23	0.90
0.39	0.84	0.65	0.90	0.19	0.95
0.40	0.84	0.65	0.91	0.19	0.95
0.41	0.84	0.70	0.92	0.19	0.95
0.42	0.84	0.70	0.93	0.19	0.95
0.43	0.84	0.70	0.94	0.16	0.95
0.44	0.84	0.70	0.95	0.14	0.95
0.45	0.77	0.70	0.96	0.14	0.95
0.46	0.77	0.70	0.97	0.14	0.95
0.47	0.77	0.70	0.98	0.14	0.95
0.48	0.77	0.70	0.99	0.14	0.95
0.49	0.77	0.70	1.00	0.14	0.95
0.50	0.77	0.70			

Ages 3.06–3.11

Score	Specificity n = 50	Sensitivity n = 24	Score	Specificity n = 50	Sensitivity n = 24
0.00	1.00	0.00	0.51	0.92	0.71
0.01	0.98	0.25	0.52	0.92	0.71
0.02	0.98	0.25	0.53	0.92	0.71
0.03	0.98	0.25	0.54	0.88	0.71
0.04	0.98	0.25	0.55	0.88	0.71
0.05	0.98	0.25	0.56	0.88	0.75
0.06	0.98	0.25	0.57	0.84	0.79
0.07	0.98	0.25	0.58	0.84	0.79
0.08	0.98	0.25	0.59	0.84	0.79
0.09	0.98	0.25	0.60	0.84	0.83
0.10	0.98	0.33	0.61	0.84	0.88
0.11	0.98	0.33	0.62	0.82	0.88
0.12	0.98	0.33	0.63	0.80	0.88
0.13	0.98	0.33	0.64	0.78	0.88
0.14	0.98	0.38	0.65	0.78	0.92
0.15	0.98	0.42	0.66	0.78	0.92
0.16	0.98	0.42	0.67	0.78	0.92
0.17	0.98	0.42	0.68	0.72	0.92
0.18	0.98	0.42	0.69	0.72	0.92
0.19	0.98	0.42	0.70	0.68	0.96
0.20	0.96	0.42	0.71	0.68	0.96
0.21	0.96	0.42	0.72	0.66	1.00
0.22	0.96	0.42	0.73	0.62	1.00
0.23	0.96	0.46	0.74	0.62	1.00
0.24	0.96	0.46	0.75	0.62	1.00
0.25	0.96	0.46	0.76	0.60	1.00
0.26	0.94	0.50	0.77	0.58	1.00
0.27	0.94	0.50	0.78	0.58	1.00
0.28	0.94	0.50	0.79	0.56	1.00
0.29	0.94	0.50	0.80	0.52	1.00
0.30	0.94	0.50	0.81	0.50	1.00
0.31	0.94	0.50	0.82	0.48	1.00
0.32	0.94	0.50	0.83	0.46	1.00
0.33	0.94	0.50	0.84	0.44	1.00
0.34	0.94	0.63	0.85	0.44	1.00
0.35	0.94	0.63	0.86	0.44	1.00
0.36	0.94	0.63	0.87	0.44	1.00
0.37	0.94	0.63	0.88	0.40	1.00
0.38	0.94	0.63	0.89	0.40	1.00
0.39	0.94	0.67	0.90	0.36	1.00
0.40	0.94	0.67	0.91	0.34	1.00
0.41	0.94	0.67	0.92	0.34	1.00
0.42	0.94	0.67	0.93	0.30	1.00
0.43	0.94	0.67	0.94	0.28	1.00
0.44	0.94	0.67	0.95	0.26	1.00
0.45	0.94	0.67	0.96	0.26	1.00
0.46	0.94	0.67	0.97	0.26	1.00
0.47	0.94	0.67	0.98	0.26	1.00
0.48	0.92	0.67	0.99	0.26	1.00
0.49	0.92	0.67	1.00	0.26	1.00
0.50	0.92	0.67			

Ages 4.00–4.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	0.92	0.64
0.01	1.00	0.18	0.52	0.92	0.64
0.02	1.00	0.18	0.53	0.92	0.64
0.03	1.00	0.18	0.54	0.92	0.66
0.04	1.00	0.18	0.55	0.92	0.66
0.05	1.00	0.18	0.56	0.92	0.66
0.06	1.00	0.18	0.57	0.92	0.66
0.07	1.00	0.20	0.58	0.92	0.70
0.08	1.00	0.20	0.59	0.92	0.70
0.09	1.00	0.22	0.60	0.92	0.70
0.10	1.00	0.22	0.61	0.92	0.70
0.11	1.00	0.24	0.62	0.90	0.76
0.12	1.00	0.26	0.63	0.88	0.76
0.13	1.00	0.26	0.64	0.88	0.76
0.14	0.98	0.26	0.65	0.88	0.76
0.15	0.98	0.26	0.66	0.86	0.76
0.16	0.98	0.30	0.67	0.86	0.76
0.17	0.98	0.30	0.68	0.84	0.84
0.18	0.98	0.32	0.69	0.84	0.84
0.19	0.98	0.36	0.70	0.82	0.84
0.20	0.98	0.38	0.71	0.82	0.86
0.21	0.98	0.38	0.72	0.82	0.88
0.22	0.98	0.38	0.73	0.80	0.88
0.23	0.98	0.38	0.74	0.80	0.90
0.24	0.98	0.38	0.75	0.80	0.90
0.25	0.98	0.38	0.76	0.76	0.90
0.26	0.98	0.42	0.77	0.70	0.90
0.27	0.98	0.42	0.78	0.70	0.90
0.28	0.98	0.42	0.79	0.66	0.90
0.29	0.98	0.42	0.80	0.66	0.90
0.30	0.98	0.44	0.81	0.66	0.90
0.31	0.98	0.46	0.82	0.66	0.92
0.32	0.98	0.46	0.83	0.62	0.92
0.33	0.98	0.46	0.84	0.60	0.92
0.34	0.98	0.54	0.85	0.60	0.92
0.35	0.98	0.54	0.86	0.60	0.92
0.36	0.98	0.54	0.87	0.58	0.92
0.37	0.98	0.54	0.88	0.58	0.92
0.38	0.98	0.54	0.89	0.56	0.94
0.39	0.98	0.58	0.90	0.50	0.94
0.40	0.96	0.60	0.91	0.50	0.94
0.41	0.94	0.60	0.92	0.50	0.94
0.42	0.92	0.60	0.93	0.50	0.94
0.43	0.92	0.60	0.94	0.48	0.94
0.44	0.92	0.60	0.95	0.34	0.96
0.45	0.92	0.60	0.96	0.34	0.96
0.46	0.92	0.60	0.97	0.34	0.96
0.47	0.92	0.60	0.98	0.34	0.96
0.48	0.92	0.62	0.99	0.34	0.96
0.49	0.92	0.62	1.00	0.34	0.96
0.50	0.92	0.62			

Ages 4.06–4.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.56
0.01	1.00	0.06	0.52	1.00	0.56
0.02	1.00	0.06	0.53	1.00	0.56
0.03	1.00	0.06	0.54	1.00	0.58
0.04	1.00	0.06	0.55	1.00	0.62
0.05	1.00	0.06	0.56	1.00	0.62
0.06	1.00	0.06	0.57	1.00	0.64
0.07	1.00	0.08	0.58	1.00	0.64
0.08	1.00	0.08	0.59	1.00	0.64
0.09	1.00	0.08	0.60	1.00	0.64
0.10	1.00	0.10	0.61	1.00	0.68
0.11	1.00	0.10	0.62	1.00	0.68
0.12	1.00	0.12	0.63	1.00	0.70
0.13	1.00	0.12	0.64	1.00	0.70
0.14	1.00	0.12	0.65	0.98	0.72
0.15	1.00	0.12	0.66	0.98	0.72
0.16	1.00	0.16	0.67	0.98	0.72
0.17	1.00	0.16	0.68	0.96	0.74
0.18	1.00	0.16	0.69	0.96	0.74
0.19	1.00	0.16	0.70	0.94	0.74
0.20	1.00	0.18	0.71	0.94	0.78
0.21	1.00	0.20	0.72	0.94	0.78
0.22	1.00	0.22	0.73	0.92	0.82
0.23	1.00	0.22	0.74	0.92	0.82
0.24	1.00	0.22	0.75	0.92	0.82
0.25	1.00	0.22	0.76	0.92	0.86
0.26	1.00	0.22	0.77	0.92	0.86
0.27	1.00	0.22	0.78	0.92	0.86
0.28	1.00	0.24	0.79	0.86	0.86
0.29	1.00	0.26	0.80	0.86	0.86
0.30	1.00	0.26	0.81	0.86	0.86
0.31	1.00	0.28	0.82	0.84	0.86
0.32	1.00	0.28	0.83	0.78	0.92
0.33	1.00	0.28	0.84	0.66	0.92
0.34	1.00	0.32	0.85	0.66	0.92
0.35	1.00	0.32	0.86	0.66	0.92
0.36	1.00	0.32	0.87	0.66	0.92
0.37	1.00	0.34	0.88	0.64	0.92
0.38	1.00	0.34	0.89	0.62	0.92
0.39	1.00	0.38	0.90	0.52	0.92
0.40	1.00	0.38	0.91	0.52	0.92
0.41	1.00	0.42	0.92	0.50	0.92
0.42	1.00	0.42	0.93	0.50	0.92
0.43	1.00	0.42	0.94	0.46	0.94
0.44	1.00	0.42	0.95	0.32	0.98
0.45	1.00	0.46	0.96	0.32	0.98
0.46	1.00	0.46	0.97	0.32	0.98
0.47	1.00	0.50	0.98	0.32	0.98
0.48	1.00	0.50	0.99	0.32	0.98
0.49	1.00	0.50	1.00	0.32	0.98
0.50	1.00	0.50			

Ages 5.00–5.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.54
0.01	1.00	0.10	0.52	1.00	0.54
0.02	1.00	0.10	0.53	1.00	0.54
0.03	1.00	0.10	0.54	1.00	0.56
0.04	1.00	0.10	0.55	1.00	0.56
0.05	1.00	0.10	0.56	1.00	0.56
0.06	1.00	0.10	0.57	1.00	0.64
0.07	1.00	0.12	0.58	1.00	0.64
0.08	1.00	0.16	0.59	0.98	0.64
0.09	1.00	0.18	0.60	0.98	0.64
0.10	1.00	0.18	0.61	0.98	0.64
0.11	1.00	0.20	0.62	0.96	0.64
0.12	1.00	0.24	0.63	0.96	0.64
0.13	1.00	0.26	0.64	0.96	0.66
0.14	1.00	0.28	0.65	0.96	0.70
0.15	1.00	0.28	0.66	0.94	0.70
0.16	1.00	0.30	0.67	0.94	0.70
0.17	1.00	0.30	0.68	0.90	0.72
0.18	1.00	0.30	0.69	0.90	0.72
0.19	1.00	0.32	0.70	0.90	0.72
0.20	1.00	0.32	0.71	0.90	0.72
0.21	1.00	0.32	0.72	0.90	0.74
0.22	1.00	0.32	0.73	0.86	0.80
0.23	1.00	0.38	0.74	0.86	0.80
0.24	1.00	0.38	0.75	0.86	0.80
0.25	1.00	0.38	0.76	0.84	0.80
0.26	1.00	0.40	0.77	0.78	0.80
0.27	1.00	0.40	0.78	0.78	0.80
0.28	1.00	0.40	0.79	0.76	0.88
0.29	1.00	0.42	0.80	0.74	0.88
0.30	1.00	0.42	0.81	0.74	0.90
0.31	1.00	0.42	0.82	0.74	0.90
0.32	1.00	0.42	0.83	0.72	0.90
0.33	1.00	0.42	0.84	0.68	0.92
0.34	1.00	0.42	0.85	0.68	0.92
0.35	1.00	0.42	0.86	0.68	0.92
0.36	1.00	0.42	0.87	0.66	0.92
0.37	1.00	0.42	0.88	0.66	0.92
0.38	1.00	0.42	0.89	0.58	0.92
0.39	1.00	0.42	0.90	0.54	0.92
0.40	1.00	0.44	0.91	0.54	0.92
0.41	1.00	0.46	0.92	0.54	0.92
0.42	1.00	0.46	0.93	0.54	0.94
0.43	1.00	0.46	0.94	0.54	0.96
0.44	1.00	0.46	0.95	0.32	0.96
0.45	1.00	0.48	0.96	0.32	0.96
0.46	1.00	0.48	0.97	0.32	0.96
0.47	1.00	0.48	0.98	0.32	0.96
0.48	1.00	0.50	0.99	0.32	0.96
0.49	1.00	0.50	1.00	0.32	0.96
0.50	1.00	0.50			

Ages 5.06–5.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity| n = 50
0.00	1.00	0.00	0.51	1.00	0.52
0.01	1.00	0.06	0.52	1.00	0.52
0.02	1.00	0.06	0.53	1.00	0.52
0.03	1.00	0.06	0.54	1.00	0.54
0.04	1.00	0.06	0.55	1.00	0.54
0.05	1.00	0.06	0.56	1.00	0.54
0.06	1.00	0.06	0.57	1.00	0.58
0.07	1.00	0.10	0.58	1.00	0.58
0.08	1.00	0.10	0.59	1.00	0.58
0.09	1.00	0.10	0.60	1.00	0.58
0.10	1.00	0.10	0.61	1.00	0.58
0.11	1.00	0.10	0.62	1.00	0.60
0.12	1.00	0.14	0.63	1.00	0.60
0.13	1.00	0.16	0.64	1.00	0.60
0.14	1.00	0.20	0.65	0.98	0.62
0.15	1.00	0.20	0.66	0.98	0.64
0.16	1.00	0.20	0.67	0.98	0.64
0.17	1.00	0.20	0.68	0.98	0.68
0.18	1.00	0.20	0.69	0.98	0.68
0.19	1.00	0.22	0.70	0.98	0.68
0.20	1.00	0.22	0.71	0.98	0.68
0.21	1.00	0.22	0.72	0.98	0.70
0.22	1.00	0.22	0.73	0.96	0.74
0.23	1.00	0.24	0.74	0.96	0.74
0.24	1.00	0.24	0.75	0.96	0.74
0.25	1.00	0.24	0.76	0.94	0.74
0.26	1.00	0.26	0.77	0.92	0.74
0.27	1.00	0.26	0.78	0.92	0.74
0.28	1.00	0.26	0.79	0.92	0.82
0.29	1.00	0.30	0.80	0.90	0.84
0.30	1.00	0.30	0.81	0.90	0.84
0.31	1.00	0.32	0.82	0.90	0.84
0.32	1.00	0.34	0.83	0.90	0.84
0.33	1.00	0.34	0.84	0.86	0.86
0.34	1.00	0.38	0.85	0.86	0.86
0.35	1.00	0.38	0.86	0.86	0.86
0.36	1.00	0.38	0.87	0.86	0.86
0.37	1.00	0.40	0.88	0.86	0.86
0.38	1.00	0.40	0.89	0.82	0.88
0.39	1.00	0.40	0.90	0.72	0.94
0.40	1.00	0.40	0.91	0.72	0.94
0.41	1.00	0.40	0.92	0.72	0.94
0.42	1.00	0.40	0.93	0.72	0.94
0.43	1.00	0.40	0.94	0.68	0.94
0.44	1.00	0.40	0.95	0.44	1.00
0.45	1.00	0.40	0.96	0.44	1.00
0.46	1.00	0.40	0.97	0.44	1.00
0.47	1.00	0.40	0.98	0.44	1.00
0.48	1.00	0.42	0.99	0.44	1.00
0.49	1.00	0.42	1.00	0.44	1.00
0.50	1.00	0.42			

Ages 6.00–6.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.40
0.01	1.00	0.00	0.52	1.00	0.40
0.02	1.00	0.00	0.53	1.00	0.40
0.03	1.00	0.00	0.54	1.00	0.42
0.04	1.00	0.00	0.55	1.00	0.42
0.05	1.00	0.00	0.56	1.00	0.42
0.06	1.00	0.00	0.57	1.00	0.50
0.07	1.00	0.02	0.58	1.00	0.50
0.08	1.00	0.02	0.59	1.00	0.50
0.09	1.00	0.02	0.60	1.00	0.50
0.10	1.00	0.02	0.61	1.00	0.50
0.11	1.00	0.02	0.62	1.00	0.52
0.12	1.00	0.02	0.63	1.00	0.52
0.13	1.00	0.02	0.64	1.00	0.54
0.14	1.00	0.02	0.65	1.00	0.54
0.15	1.00	0.02	0.66	1.00	0.56
0.16	1.00	0.02	0.67	1.00	0.56
0.17	1.00	0.02	0.68	0.96	0.58
0.18	1.00	0.02	0.69	0.96	0.58
0.19	1.00	0.02	0.70	0.96	0.58
0.20	1.00	0.02	0.71	0.96	0.58
0.21	1.00	0.04	0.72	0.96	0.62
0.22	1.00	0.04	0.73	0.92	0.64
0.23	1.00	0.06	0.74	0.92	0.64
0.24	1.00	0.06	0.75	0.92	0.64
0.25	1.00	0.06	0.76	0.92	0.66
0.26	1.00	0.08	0.77	0.92	0.66
0.27	1.00	0.08	0.78	0.92	0.66
0.28	1.00	0.08	0.79	0.92	0.78
0.29	1.00	0.08	0.80	0.92	0.78
0.30	1.00	0.10	0.81	0.92	0.80
0.31	1.00	0.10	0.82	0.92	0.80
0.32	1.00	0.14	0.83	0.88	0.80
0.33	1.00	0.14	0.84	0.86	0.80
0.34	1.00	0.18	0.85	0.86	0.80
0.35	1.00	0.18	0.86	0.86	0.82
0.36	1.00	0.18	0.87	0.86	0.84
0.37	1.00	0.18	0.88	0.86	0.84
0.38	1.00	0.18	0.89	0.86	0.86
0.39	1.00	0.18	0.90	0.80	0.92
0.40	1.00	0.22	0.91	0.80	0.92
0.41	1.00	0.24	0.92	0.80	0.92
0.42	1.00	0.28	0.93	0.78	0.92
0.43	1.00	0.28	0.94	0.78	0.92
0.44	1.00	0.28	0.95	0.34	0.94
0.45	1.00	0.32	0.96	0.34	0.94
0.46	1.00	0.32	0.97	0.34	0.94
0.47	1.00	0.32	0.98	0.34	0.94
0.48	1.00	0.34	0.99	0.34	0.94
0.49	1.00	0.34	1.00	0.34	0.94
0.50	1.00	0.34			

Ages 6.06–6.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.44
0.01	1.00	0.02	0.52	1.00	0.44
0.02	1.00	0.02	0.53	1.00	0.44
0.03	1.00	0.02	0.54	1.00	0.44
0.04	1.00	0.02	0.55	1.00	0.44
0.05	1.00	0.02	0.56	1.00	0.44
0.06	1.00	0.02	0.57	1.00	0.50
0.07	1.00	0.02	0.58	1.00	0.50
0.08	1.00	0.02	0.59	1.00	0.50
0.09	1.00	0.02	0.60	1.00	0.50
0.10	1.00	0.02	0.61	1.00	0.50
0.11	1.00	0.02	0.62	1.00	0.54
0.12	1.00	0.04	0.63	1.00	0.54
0.13	1.00	0.04	0.64	1.00	0.54
0.14	1.00	0.06	0.65	1.00	0.54
0.15	1.00	0.06	0.66	1.00	0.54
0.16	1.00	0.06	0.67	1.00	0.54
0.17	1.00	0.06	0.68	1.00	0.58
0.18	1.00	0.08	0.69	1.00	0.58
0.19	1.00	0.10	0.70	1.00	0.58
0.20	1.00	0.10	0.71	1.00	0.58
0.21	1.00	0.10	0.72	1.00	0.60
0.22	1.00	0.10	0.73	1.00	0.62
0.23	1.00	0.10	0.74	1.00	0.62
0.24	1.00	0.10	0.75	1.00	0.62
0.25	1.00	0.14	0.76	1.00	0.62
0.26	1.00	0.14	0.77	1.00	0.62
0.27	1.00	0.14	0.78	1.00	0.64
0.28	1.00	0.14	0.79	0.98	0.70
0.29	1.00	0.20	0.80	0.98	0.70
0.30	1.00	0.20	0.81	0.98	0.70
0.31	1.00	0.20	0.82	0.98	0.74
0.32	1.00	0.22	0.83	0.98	0.76
0.33	1.00	0.22	0.84	0.88	0.78
0.34	1.00	0.28	0.85	0.88	0.78
0.35	1.00	0.28	0.86	0.88	0.78
0.36	1.00	0.28	0.87	0.88	0.80
0.37	1.00	0.30	0.88	0.88	0.80
0.38	1.00	0.30	0.89	0.88	0.84
0.39	1.00	0.30	0.90	0.74	0.86
0.40	1.00	0.34	0.91	0.74	0.86
0.41	1.00	0.36	0.92	0.74	0.86
0.42	1.00	0.36	0.93	0.72	0.86
0.43	1.00	0.36	0.94	0.70	0.86
0.44	1.00	0.36	0.95	0.44	0.92
0.45	1.00	0.36	0.96	0.44	0.92
0.46	1.00	0.36	0.97	0.44	0.92
0.47	1.00	0.36	0.98	0.44	0.92
0.48	1.00	0.38	0.99	0.44	0.92
0.49	1.00	0.38	1.00	0.44	0.92
0.50	1.00	0.38			

Ages 7.00–7.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.02	0.51	0.14
0.01	0.02	0.52	0.14
0.02	0.02	0.53	0.18
0.03	0.02	0.54	0.18
0.04	0.02	0.55	0.18
0.05	0.02	0.56	0.20
0.06	0.02	0.57	0.20
0.07	0.02	0.58	0.20
0.08	0.02	0.59	0.20
0.09	0.02	0.60	0.22
0.10	0.02	0.61	0.24
0.11	0.02	0.62	0.24
0.12	0.02	0.63	0.26
0.13	0.02	0.64	0.28
0.14	0.02	0.65	0.28
0.15	0.02	0.66	0.28
0.16	0.02	0.67	0.30
0.17	0.02	0.68	0.30
0.18	0.02	0.69	0.30
0.19	0.02	0.70	0.30
0.20	0.02	0.71	0.36
0.21	0.02	0.72	0.38
0.22	0.02	0.73	0.38
0.23	0.02	0.74	0.38
0.24	0.04	0.75	0.38
0.25	0.04	0.76	0.40
0.26	0.04	0.77	0.40
0.27	0.04	0.78	0.48
0.28	0.04	0.79	0.48
0.29	0.06	0.80	0.50
0.30	0.06	0.81	0.50
0.31	0.06	0.82	0.52
0.32	0.06	0.83	0.54
0.33	0.08	0.84	0.54
0.34	0.08	0.85	0.54
0.35	0.08	0.86	0.54
0.36	0.08	0.87	0.54
0.37	0.08	0.88	0.60
0.38	0.10	0.89	0.66
0.39	0.12	0.90	0.68
0.40	0.12	0.91	0.68
0.41	0.12	0.92	0.68
0.42	0.12	0.93	0.68
0.43	0.12	0.94	0.84
0.44	0.12	0.95	0.84
0.45	0.12	0.96	0.84
0.46	0.12	0.97	0.84
0.47	0.12	0.98	0.84
0.48	0.12	0.99	0.84
0.49	0.12	1.00	1.00
0.50	0.14		

Ages 8.00–8.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.00	0.51	0.16
0.01	0.00	0.52	0.16
0.02	0.00	0.53	0.18
0.03	0.00	0.54	0.18
0.04	0.00	0.55	0.18
0.05	0.00	0.56	0.20
0.06	0.00	0.57	0.20
0.07	0.00	0.58	0.20
0.08	0.00	0.59	0.20
0.09	0.00	0.60	0.20
0.10	0.00	0.61	0.22
0.11	0.00	0.62	0.22
0.12	0.00	0.63	0.22
0.13	0.00	0.64	0.22
0.14	0.00	0.65	0.22
0.15	0.00	0.66	0.22
0.16	0.00	0.67	0.24
0.17	0.02	0.68	0.24
0.18	0.02	0.69	0.24
0.19	0.02	0.70	0.24
0.20	0.02	0.71	0.24
0.21	0.02	0.72	0.28
0.22	0.04	0.73	0.28
0.23	0.04	0.74	0.28
0.24	0.04	0.75	0.32
0.25	0.06	0.76	0.32
0.26	0.06	0.77	0.32
0.27	0.06	0.78	0.38
0.28	0.08	0.79	0.38
0.29	0.08	0.80	0.40
0.30	0.08	0.81	0.40
0.31	0.10	0.82	0.40
0.32	0.10	0.83	0.50
0.33	0.12	0.84	0.50
0.34	0.12	0.85	0.50
0.35	0.12	0.86	0.50
0.36	0.12	0.87	0.50
0.37	0.12	0.88	0.52
0.38	0.12	0.89	0.60
0.39	0.14	0.90	0.60
0.40	0.14	0.91	0.60
0.41	0.14	0.92	0.60
0.42	0.14	0.93	0.64
0.43	0.14	0.94	0.80
0.44	0.16	0.95	0.80
0.45	0.16	0.96	0.80
0.46	0.16	0.97	0.80
0.47	0.16	0.98	0.80
0.48	0.16	0.99	0.80
0.49	0.16	1.00	1.00
0.50	0.16		

Ages 3.00–3.05

Score	Specificity n = 43	Sensitivity n = 20	Score	Specificity n = 43	Sensitivity n = 20
0.00	1.00	0.00	0.51	0.77	0.90
0.01	1.00	0.15	0.52	0.77	0.90
0.02	1.00	0.15	0.53	0.77	0.90
0.03	1.00	0.15	0.54	0.77	0.90
0.04	1.00	0.15	0.55	0.74	0.90
0.05	1.00	0.15	0.56	0.74	0.90
0.06	1.00	0.15	0.57	0.74	0.90
0.07	1.00	0.15	0.58	0.74	0.90
0.08	1.00	0.15	0.59	0.74	0.90
0.09	1.00	0.15	0.60	0.70	0.90
0.10	1.00	0.20	0.61	0.70	0.90
0.11	1.00	0.20	0.62	0.67	0.90
0.12	1.00	0.20	0.63	0.65	0.90
0.13	1.00	0.20	0.64	0.63	0.90
0.14	1.00	0.20	0.65	0.60	0.90
0.15	1.00	0.20	0.66	0.60	0.90
0.16	1.00	0.25	0.67	0.60	0.90
0.17	1.00	0.25	0.68	0.60	0.90
0.18	1.00	0.30	0.69	0.56	0.90
0.19	1.00	0.35	0.70	0.56	0.90
0.20	1.00	0.35	0.71	0.51	0.90
0.21	1.00	0.40	0.72	0.47	0.90
0.22	1.00	0.40	0.73	0.42	0.90
0.23	1.00	0.40	0.74	0.42	0.90
0.24	1.00	0.40	0.75	0.42	0.90
0.25	1.00	0.40	0.76	0.42	0.90
0.26	0.95	0.40	0.77	0.40	0.90
0.27	0.95	0.40	0.78	0.40	0.90
0.28	0.95	0.40	0.79	0.40	0.90
0.29	0.93	0.40	0.80	0.40	0.90
0.30	0.91	0.45	0.81	0.37	0.90
0.31	0.91	0.50	0.82	0.33	0.90
0.32	0.91	0.55	0.83	0.30	0.90
0.33	0.88	0.55	0.84	0.30	0.90
0.34	0.88	0.55	0.85	0.30	0.90
0.35	0.88	0.55	0.86	0.28	0.95
0.36	0.86	0.55	0.87	0.28	0.95
0.37	0.86	0.60	0.88	0.26	1.00
0.38	0.86	0.65	0.89	0.26	1.00
0.39	0.86	0.65	0.90	0.19	1.00
0.40	0.86	0.65	0.91	0.14	1.00
0.41	0.86	0.65	0.92	0.14	1.00
0.42	0.86	0.65	0.93	0.12	1.00
0.43	0.86	0.65	0.94	0.12	1.00
0.44	0.86	0.65	0.95	0.12	1.00
0.45	0.86	0.65	0.96	0.07	1.00
0.46	0.86	0.65	0.97	0.07	1.00
0.47	0.86	0.80	0.98	0.05	1.00
0.48	0.84	0.80	0.99	0.05	1.00
0.49	0.81	0.80	1.00	0.05	1.00
0.50	0.81	0.80			

Ages 3.06–3.11

Score	Specificity n = 50	Sensitivity n = 24	Score	Specificity n = 50	Sensitivity n = 24
0.00	1.00	0.00	0.51	0.86	0.79
0.01	1.00	0.21	0.52	0.86	0.79
0.02	1.00	0.21	0.53	0.84	0.79
0.03	1.00	0.21	0.54	0.82	0.79
0.04	1.00	0.21	0.55	0.80	0.79
0.05	1.00	0.21	0.56	0.80	0.79
0.06	1.00	0.21	0.57	0.80	0.79
0.07	1.00	0.21	0.58	0.80	0.79
0.08	1.00	0.25	0.59	0.80	0.79
0.09	1.00	0.25	0.60	0.80	0.79
0.10	1.00	0.25	0.61	0.80	0.79
0.11	1.00	0.29	0.62	0.80	0.83
0.12	1.00	0.33	0.63	0.78	0.83
0.13	1.00	0.33	0.64	0.78	0.83
0.14	1.00	0.42	0.65	0.78	0.83
0.15	1.00	0.42	0.66	0.78	0.83
0.16	1.00	0.42	0.67	0.78	0.83
0.17	1.00	0.42	0.68	0.76	0.83
0.18	1.00	0.46	0.69	0.76	0.83
0.19	1.00	0.54	0.70	0.76	0.83
0.20	1.00	0.58	0.71	0.76	0.83
0.21	0.98	0.58	0.72	0.72	0.83
0.22	0.98	0.58	0.73	0.72	0.83
0.23	0.98	0.58	0.74	0.70	0.88
0.24	0.98	0.58	0.75	0.70	0.88
0.25	0.98	0.58	0.76	0.64	0.92
0.26	0.98	0.58	0.77	0.62	0.92
0.27	0.98	0.58	0.78	0.60	0.92
0.28	0.98	0.63	0.79	0.58	0.92
0.29	0.98	0.63	0.80	0.58	0.92
0.30	0.98	0.63	0.81	0.58	0.96
0.31	0.98	0.63	0.82	0.58	0.96
0.32	0.98	0.63	0.83	0.52	0.96
0.33	0.98	0.67	0.84	0.52	0.96
0.34	0.98	0.71	0.85	0.52	0.96
0.35	0.98	0.71	0.86	0.48	1.00
0.36	0.96	0.71	0.87	0.42	1.00
0.37	0.96	0.71	0.88	0.40	1.00
0.38	0.96	0.71	0.89	0.40	1.00
0.39	0.96	0.71	0.90	0.40	1.00
0.40	0.96	0.75	0.91	0.38	1.00
0.41	0.94	0.79	0.92	0.36	1.00
0.42	0.94	0.79	0.93	0.36	1.00
0.43	0.92	0.79	0.94	0.36	1.00
0.44	0.92	0.79	0.95	0.34	1.00
0.45	0.92	0.79	0.96	0.28	1.00
0.46	0.92	0.79	0.97	0.24	1.00
0.47	0.90	0.79	0.98	0.22	1.00
0.48	0.88	0.79	0.99	0.22	1.00
0.49	0.88	0.79	1.00	0.22	1.00
0.50	0.88	0.79			

Ages 4.00–4.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	0.92	0.66
0.01	1.00	0.10	0.52	0.92	0.70
0.02	1.00	0.10	0.53	0.92	0.70
0.03	1.00	0.10	0.54	0.92	0.70
0.04	1.00	0.10	0.55	0.92	0.72
0.05	1.00	0.10	0.56	0.92	0.74
0.06	1.00	0.12	0.57	0.90	0.74
0.07	1.00	0.12	0.58	0.90	0.74
0.08	0.98	0.12	0.59	0.90	0.74
0.09	0.98	0.14	0.60	0.90	0.76
0.10	0.98	0.16	0.61	0.90	0.76
0.11	0.98	0.18	0.62	0.90	0.76
0.12	0.98	0.18	0.63	0.90	0.80
0.13	0.98	0.18	0.64	0.90	0.80
0.14	0.98	0.22	0.65	0.90	0.82
0.15	0.98	0.22	0.66	0.90	0.84
0.16	0.98	0.24	0.67	0.88	0.84
0.17	0.98	0.28	0.68	0.88	0.86
0.18	0.98	0.28	0.69	0.86	0.86
0.19	0.98	0.28	0.70	0.86	0.88
0.20	0.98	0.32	0.71	0.86	0.88
0.21	0.96	0.32	0.72	0.84	0.88
0.22	0.96	0.32	0.73	0.84	0.88
0.23	0.96	0.34	0.74	0.84	0.88
0.24	0.96	0.36	0.75	0.84	0.88
0.25	0.96	0.38	0.76	0.82	0.88
0.26	0.96	0.40	0.77	0.82	0.88
0.27	0.96	0.40	0.78	0.78	0.90
0.28	0.96	0.40	0.79	0.78	0.90
0.29	0.96	0.40	0.80	0.78	0.92
0.30	0.96	0.46	0.81	0.78	0.92
0.31	0.96	0.46	0.82	0.72	0.92
0.32	0.96	0.48	0.83	0.72	0.92
0.33	0.96	0.50	0.84	0.70	0.92
0.34	0.96	0.54	0.85	0.68	0.92
0.35	0.96	0.56	0.86	0.68	0.92
0.36	0.96	0.58	0.87	0.66	0.92
0.37	0.96	0.58	0.88	0.66	0.92
0.38	0.96	0.58	0.89	0.62	0.92
0.39	0.96	0.58	0.90	0.60	0.92
0.40	0.96	0.58	0.91	0.60	0.92
0.41	0.96	0.58	0.92	0.54	0.94
0.42	0.96	0.58	0.93	0.52	0.94
0.43	0.96	0.58	0.94	0.52	0.94
0.44	0.96	0.60	0.95	0.48	0.94
0.45	0.94	0.60	0.96	0.38	0.94
0.46	0.92	0.60	0.97	0.38	0.94
0.47	0.92	0.60	0.98	0.26	0.96
0.48	0.92	0.62	0.99	0.26	0.96
0.49	0.92	0.64	1.00	0.26	0.96
0.50	0.92	0.66			

Ages 4.06–4.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	0.98	0.68
0.01	1.00	0.06	0.52	0.98	0.68
0.02	1.00	0.06	0.53	0.98	0.68
0.03	1.00	0.06	0.54	0.98	0.70
0.04	1.00	0.06	0.55	0.98	0.70
0.05	1.00	0.06	0.56	0.98	0.74
0.06	1.00	0.08	0.57	0.98	0.74
0.07	1.00	0.08	0.58	0.98	0.76
0.08	1.00	0.08	0.59	0.98	0.76
0.09	1.00	0.12	0.60	0.98	0.78
0.10	1.00	0.12	0.61	0.98	0.78
0.11	1.00	0.14	0.62	0.96	0.78
0.12	1.00	0.14	0.63	0.96	0.78
0.13	1.00	0.14	0.64	0.96	0.78
0.14	1.00	0.14	0.65	0.94	0.80
0.15	1.00	0.14	0.66	0.94	0.80
0.16	1.00	0.16	0.67	0.94	0.82
0.17	1.00	0.18	0.68	0.94	0.82
0.18	1.00	0.18	0.69	0.94	0.82
0.19	1.00	0.18	0.70	0.94	0.82
0.20	1.00	0.18	0.71	0.94	0.82
0.21	1.00	0.18	0.72	0.94	0.82
0.22	1.00	0.18	0.73	0.94	0.82
0.23	1.00	0.18	0.74	0.94	0.84
0.24	1.00	0.18	0.75	0.94	0.84
0.25	1.00	0.18	0.76	0.94	0.86
0.26	1.00	0.18	0.77	0.94	0.88
0.27	1.00	0.18	0.78	0.94	0.88
0.28	1.00	0.24	0.79	0.90	0.88
0.29	1.00	0.24	0.80	0.88	0.90
0.30	1.00	0.24	0.81	0.88	0.90
0.31	1.00	0.26	0.82	0.88	0.90
0.32	1.00	0.30	0.83	0.86	0.90
0.33	1.00	0.32	0.84	0.86	0.92
0.34	1.00	0.34	0.85	0.84	0.92
0.35	1.00	0.36	0.86	0.82	0.92
0.36	1.00	0.38	0.87	0.78	0.92
0.37	1.00	0.38	0.88	0.74	0.92
0.38	1.00	0.38	0.89	0.72	0.92
0.39	1.00	0.44	0.90	0.66	0.92
0.40	1.00	0.44	0.91	0.58	0.92
0.41	1.00	0.44	0.92	0.54	0.94
0.42	1.00	0.50	0.93	0.40	0.98
0.43	1.00	0.52	0.94	0.40	0.98
0.44	1.00	0.54	0.95	0.40	0.98
0.45	1.00	0.58	0.96	0.34	0.98
0.46	1.00	0.62	0.97	0.34	0.98
0.47	1.00	0.62	0.98	0.28	1.00
0.48	0.98	0.64	0.99	0.28	1.00
0.49	0.98	0.66	1.00	0.28	1.00
0.50	0.98	0.66			

Ages 5.00–5.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.58
0.01	1.00	0.02	0.52	1.00	0.62
0.02	1.00	0.02	0.53	1.00	0.62
0.03	1.00	0.02	0.54	1.00	0.66
0.04	1.00	0.04	0.55	1.00	0.68
0.05	1.00	0.04	0.56	1.00	0.68
0.06	1.00	0.04	0.57	1.00	0.68
0.07	1.00	0.08	0.58	1.00	0.68
0.08	1.00	0.08	0.59	1.00	0.68
0.09	1.00	0.08	0.60	1.00	0.70
0.10	1.00	0.10	0.61	1.00	0.70
0.11	1.00	0.10	0.62	1.00	0.70
0.12	1.00	0.12	0.63	1.00	0.70
0.13	1.00	0.12	0.64	0.98	0.70
0.14	1.00	0.12	0.65	0.98	0.70
0.15	1.00	0.16	0.66	0.98	0.70
0.16	1.00	0.18	0.67	0.94	0.70
0.17	1.00	0.20	0.68	0.94	0.70
0.18	1.00	0.24	0.69	0.94	0.72
0.19	1.00	0.24	0.70	0.94	0.72
0.20	1.00	0.24	0.71	0.94	0.72
0.21	1.00	0.26	0.72	0.92	0.74
0.22	1.00	0.26	0.73	0.92	0.76
0.23	1.00	0.28	0.74	0.90	0.78
0.24	1.00	0.28	0.75	0.88	0.78
0.25	1.00	0.28	0.76	0.88	0.78
0.26	1.00	0.32	0.77	0.88	0.78
0.27	1.00	0.32	0.78	0.86	0.80
0.28	1.00	0.32	0.79	0.84	0.82
0.29	1.00	0.34	0.80	0.82	0.82
0.30	1.00	0.40	0.81	0.82	0.84
0.31	1.00	0.40	0.82	0.78	0.84
0.32	1.00	0.42	0.83	0.78	0.84
0.33	1.00	0.44	0.84	0.78	0.84
0.34	1.00	0.44	0.85	0.78	0.88
0.35	1.00	0.44	0.86	0.76	0.90
0.36	1.00	0.44	0.87	0.76	0.90
0.37	1.00	0.44	0.88	0.76	0.90
0.38	1.00	0.44	0.89	0.72	0.90
0.39	1.00	0.46	0.90	0.68	0.92
0.40	1.00	0.48	0.91	0.66	0.92
0.41	1.00	0.48	0.92	0.64	0.92
0.42	1.00	0.50	0.93	0.56	0.94
0.43	1.00	0.50	0.94	0.56	0.94
0.44	1.00	0.52	0.95	0.52	0.94
0.45	1.00	0.52	0.96	0.46	0.94
0.46	1.00	0.54	0.97	0.46	0.94
0.47	1.00	0.54	0.98	0.30	0.96
0.48	1.00	0.54	0.99	0.30	0.96
0.49	1.00	0.56	1.00	0.30	0.96
0.50	1.00	0.56			

Ages 5.06–5.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.52
0.01	1.00	0.04	0.52	1.00	0.52
0.02	1.00	0.04	0.53	1.00	0.52
0.03	1.00	0.04	0.54	1.00	0.52
0.04	1.00	0.06	0.55	1.00	0.54
0.05	1.00	0.06	0.56	1.00	0.54
0.06	1.00	0.06	0.57	1.00	0.56
0.07	1.00	0.08	0.58	1.00	0.56
0.08	1.00	0.10	0.59	1.00	0.56
0.09	1.00	0.10	0.60	1.00	0.58
0.10	1.00	0.12	0.61	1.00	0.60
0.11	1.00	0.14	0.62	1.00	0.62
0.12	1.00	0.18	0.63	1.00	0.64
0.13	1.00	0.18	0.64	1.00	0.64
0.14	1.00	0.20	0.65	1.00	0.66
0.15	1.00	0.20	0.66	1.00	0.70
0.16	1.00	0.20	0.67	1.00	0.70
0.17	1.00	0.20	0.68	1.00	0.72
0.18	1.00	0.22	0.69	1.00	0.72
0.19	1.00	0.24	0.70	1.00	0.72
0.20	1.00	0.24	0.71	1.00	0.72
0.21	1.00	0.24	0.72	1.00	0.74
0.22	1.00	0.24	0.73	1.00	0.74
0.23	1.00	0.26	0.74	1.00	0.74
0.24	1.00	0.26	0.75	1.00	0.74
0.25	1.00	0.26	0.76	1.00	0.76
0.26	1.00	0.26	0.77	0.98	0.78
0.27	1.00	0.30	0.78	0.96	0.78
0.28	1.00	0.30	0.79	0.96	0.78
0.29	1.00	0.30	0.80	0.94	0.80
0.30	1.00	0.32	0.81	0.94	0.80
0.31	1.00	0.32	0.82	0.94	0.82
0.32	1.00	0.32	0.83	0.94	0.84
0.33	1.00	0.32	0.84	0.92	0.84
0.34	1.00	0.32	0.85	0.90	0.88
0.35	1.00	0.34	0.86	0.90	0.88
0.36	1.00	0.34	0.87	0.90	0.90
0.37	1.00	0.36	0.88	0.90	0.90
0.38	1.00	0.36	0.89	0.88	0.90
0.39	1.00	0.38	0.90	0.88	0.94
0.40	1.00	0.40	0.91	0.84	0.94
0.41	1.00	0.42	0.92	0.84	0.94
0.42	1.00	0.42	0.93	0.76	0.96
0.43	1.00	0.44	0.94	0.76	0.96
0.44	1.00	0.46	0.95	0.72	0.96
0.45	1.00	0.46	0.96	0.62	1.00
0.46	1.00	0.48	0.97	0.62	1.00
0.47	1.00	0.48	0.98	0.38	1.00
0.48	1.00	0.50	0.99	0.38	1.00
0.49	1.00	0.50	1.00	0.38	1.00
0.50	1.00	0.50			

Ages 6.00–6.05

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.42
0.01	1.00	0.00	0.52	1.00	0.42
0.02	1.00	0.00	0.53	1.00	0.42
0.03	1.00	0.00	0.54	1.00	0.42
0.04	1.00	0.02	0.55	1.00	0.42
0.05	1.00	0.02	0.56	1.00	0.42
0.06	1.00	0.02	0.57	1.00	0.42
0.07	1.00	0.02	0.58	1.00	0.42
0.08	1.00	0.02	0.59	1.00	0.42
0.09	1.00	0.02	0.60	1.00	0.44
0.10	1.00	0.02	0.61	1.00	0.46
0.11	1.00	0.02	0.62	1.00	0.48
0.12	1.00	0.02	0.63	1.00	0.50
0.13	1.00	0.02	0.64	1.00	0.50
0.14	1.00	0.02	0.65	1.00	0.50
0.15	1.00	0.02	0.66	1.00	0.56
0.16	1.00	0.02	0.67	1.00	0.58
0.17	1.00	0.02	0.68	1.00	0.58
0.18	1.00	0.04	0.69	1.00	0.58
0.19	1.00	0.06	0.70	1.00	0.58
0.20	1.00	0.06	0.71	1.00	0.58
0.21	1.00	0.08	0.72	1.00	0.60
0.22	1.00	0.08	0.73	1.00	0.60
0.23	1.00	0.10	0.74	1.00	0.64
0.24	1.00	0.10	0.75	1.00	0.64
0.25	1.00	0.10	0.76	1.00	0.66
0.26	1.00	0.16	0.77	0.98	0.68
0.27	1.00	0.16	0.78	0.96	0.68
0.28	1.00	0.16	0.79	0.96	0.70
0.29	1.00	0.18	0.80	0.96	0.70
0.30	1.00	0.18	0.81	0.96	0.72
0.31	1.00	0.18	0.82	0.96	0.72
0.32	1.00	0.18	0.83	0.96	0.76
0.33	1.00	0.18	0.84	0.96	0.76
0.34	1.00	0.24	0.85	0.92	0.80
0.35	1.00	0.24	0.86	0.92	0.82
0.36	1.00	0.24	0.87	0.90	0.84
0.37	1.00	0.24	0.88	0.86	0.84
0.38	1.00	0.28	0.89	0.86	0.84
0.39	1.00	0.28	0.90	0.86	0.96
0.40	1.00	0.28	0.91	0.86	0.98
0.41	1.00	0.28	0.92	0.82	0.98
0.42	1.00	0.30	0.93	0.74	0.98
0.43	1.00	0.32	0.94	0.74	0.98
0.44	1.00	0.34	0.95	0.74	0.98
0.45	1.00	0.36	0.96	0.66	0.98
0.46	1.00	0.36	0.97	0.64	0.98
0.47	1.00	0.36	0.98	0.32	0.98
0.48	1.00	0.36	0.99	0.32	0.98
0.49	1.00	0.38	1.00	0.32	0.98
0.50	1.00	0.40			

Ages 6.06–6.11

Score	Specificity n = 50	Sensitivity n = 50	Score	Specificity n = 50	Sensitivity n = 50
0.00	1.00	0.00	0.51	1.00	0.38
0.01	1.00	0.00	0.52	1.00	0.38
0.02	1.00	0.00	0.53	1.00	0.40
0.03	1.00	0.00	0.54	1.00	0.42
0.04	1.00	0.00	0.55	1.00	0.44
0.05	1.00	0.00	0.56	1.00	0.46
0.06	1.00	0.00	0.57	1.00	0.50
0.07	1.00	0.02	0.58	1.00	0.50
0.08	1.00	0.04	0.59	1.00	0.52
0.09	1.00	0.04	0.60	1.00	0.54
0.10	1.00	0.06	0.61	1.00	0.54
0.11	1.00	0.06	0.62	1.00	0.54
0.12	1.00	0.06	0.63	1.00	0.54
0.13	1.00	0.06	0.64	1.00	0.54
0.14	1.00	0.06	0.65	1.00	0.54
0.15	1.00	0.06	0.66	1.00	0.54
0.16	1.00	0.06	0.67	1.00	0.56
0.17	1.00	0.06	0.68	1.00	0.56
0.18	1.00	0.08	0.69	1.00	0.60
0.19	1.00	0.08	0.70	1.00	0.62
0.20	1.00	0.10	0.71	1.00	0.62
0.21	1.00	0.14	0.72	1.00	0.64
0.22	1.00	0.14	0.73	1.00	0.66
0.23	1.00	0.16	0.74	1.00	0.68
0.24	1.00	0.16	0.75	1.00	0.70
0.25	1.00	0.16	0.76	1.00	0.70
0.26	1.00	0.18	0.77	1.00	0.70
0.27	1.00	0.20	0.78	1.00	0.70
0.28	1.00	0.22	0.79	1.00	0.72
0.29	1.00	0.22	0.80	0.98	0.74
0.30	1.00	0.22	0.81	0.98	0.74
0.31	1.00	0.22	0.82	0.98	0.74
0.32	1.00	0.24	0.83	0.98	0.74
0.33	1.00	0.24	0.84	0.98	0.74
0.34	1.00	0.24	0.85	0.96	0.76
0.35	1.00	0.24	0.86	0.94	0.78
0.36	1.00	0.24	0.87	0.92	0.78
0.37	1.00	0.24	0.88	0.90	0.80
0.38	1.00	0.24	0.89	0.90	0.82
0.39	1.00	0.24	0.90	0.86	0.84
0.40	1.00	0.26	0.91	0.82	0.84
0.41	1.00	0.28	0.92	0.80	0.88
0.42	1.00	0.28	0.93	0.68	0.92
0.43	1.00	0.28	0.94	0.68	0.92
0.44	1.00	0.28	0.95	0.68	0.92
0.45	1.00	0.28	0.96	0.54	0.94
0.46	1.00	0.32	0.97	0.54	0.94
0.47	1.00	0.32	0.98	0.36	0.98
0.48	1.00	0.34	0.99	0.36	0.98
0.49	1.00	0.36	1.00	0.36	0.98
0.50	1.00	0.38			

Ages 7.00–7.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.00	0.51	0.24
0.01	0.00	0.52	0.24
0.02	0.00	0.53	0.24
0.03	0.00	0.54	0.24
0.04	0.00	0.55	0.24
0.05	0.00	0.56	0.24
0.06	0.00	0.57	0.26
0.07	0.00	0.58	0.28
0.08	0.00	0.59	0.28
0.09	0.00	0.60	0.28
0.10	0.00	0.61	0.28
0.11	0.00	0.62	0.30
0.12	0.00	0.63	0.30
0.13	0.00	0.64	0.32
0.14	0.00	0.65	0.32
0.15	0.02	0.66	0.32
0.16	0.02	0.67	0.32
0.17	0.04	0.68	0.32
0.18	0.04	0.69	0.34
0.19	0.04	0.70	0.34
0.20	0.04	0.71	0.38
0.21	0.04	0.72	0.38
0.22	0.04	0.73	0.38
0.23	0.04	0.74	0.40
0.24	0.04	0.75	0.46
0.25	0.04	0.76	0.48
0.26	0.04	0.77	0.50
0.27	0.06	0.78	0.50
0.28	0.06	0.79	0.52
0.29	0.06	0.80	0.52
0.30	0.08	0.81	0.56
0.31	0.08	0.82	0.56
0.32	0.08	0.83	0.56
0.33	0.08	0.84	0.58
0.34	0.08	0.85	0.60
0.35	0.08	0.86	0.60
0.36	0.14	0.87	0.62
0.37	0.16	0.88	0.62
0.38	0.16	0.89	0.64
0.39	0.16	0.90	0.70
0.40	0.16	0.91	0.70
0.41	0.16	0.92	0.76
0.42	0.18	0.93	0.76
0.43	0.18	0.94	0.82
0.44	0.18	0.95	0.86
0.45	0.18	0.96	0.86
0.46	0.18	0.97	0.94
0.47	0.18	0.98	0.94
0.48	0.18	0.99	0.94
0.49	0.20	1.00	1.00
0.50	0.24		

Ages 8.00–8.11

Score	Sensitivity n = 50	Score	Sensitivity n = 50
0.00	0.00	0.51	0.18
0.01	0.00	0.52	0.18
0.02	0.00	0.53	0.20
0.03	0.00	0.54	0.22
0.04	0.00	0.55	0.24
0.05	0.00	0.56	0.26
0.06	0.00	0.57	0.28
0.07	0.00	0.58	0.28
0.08	0.00	0.59	0.28
0.09	0.02	0.60	0.28
0.10	0.02	0.61	0.28
0.11	0.02	0.62	0.32
0.12	0.02	0.63	0.32
0.13	0.02	0.64	0.32
0.14	0.04	0.65	0.32
0.15	0.04	0.66	0.32
0.16	0.04	0.67	0.34
0.17	0.04	0.68	0.34
0.18	0.04	0.69	0.34
0.19	0.06	0.70	0.34
0.20	0.06	0.71	0.34
0.21	0.06	0.72	0.34
0.22	0.06	0.73	0.34
0.23	0.06	0.74	0.36
0.24	0.06	0.75	0.36
0.25	0.06	0.76	0.36
0.26	0.06	0.77	0.38
0.27	0.06	0.78	0.42
0.28	0.06	0.79	0.44
0.29	0.06	0.80	0.44
0.30	0.08	0.81	0.48
0.31	0.08	0.82	0.48
0.32	0.08	0.83	0.48
0.33	0.10	0.84	0.48
0.34	0.10	0.85	0.50
0.35	0.10	0.86	0.50
0.36	0.10	0.87	0.50
0.37	0.10	0.88	0.50
0.38	0.12	0.89	0.52
0.39	0.12	0.90	0.54
0.40	0.14	0.91	0.54
0.41	0.14	0.92	0.64
0.42	0.14	0.93	0.64
0.43	0.16	0.94	0.66
0.44	0.16	0.95	0.68
0.45	0.16	0.96	0.68
0.46	0.18	0.97	0.84
0.47	0.18	0.98	0.84
0.48	0.18	0.99	0.84
0.49	0.18	1.00	1.00
0.50	0.18		

Probe Score Look-up Tables

C

Denominator \ Numerator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	100																																							
2	50	100																																						
3	33	67	100																																					
4	25	50	75	100																																				
5	20	40	60	80	100																																			
6	17	33	50	67	83	100																																		
7	14	29	43	57	71	86	100																																	
8	13	25	38	50	63	75	88	100																																
9	11	22	33	44	56	67	78	89	100																															
10	10	20	30	40	50	60	70	80	90	100																														
11	9	18	27	36	45	55	64	73	82	91	100																													
12	8	17	25	33	42	50	58	67	75	83	92	100																												
13	8	15	23	31	38	46	54	62	69	77	85	92	100																											
14	7	14	21	29	36	43	50	57	64	71	79	86	93	100																										
15	7	13	20	27	33	40	47	53	60	67	73	80	87	93	100																									
16	6	13	19	25	31	38	44	50	56	63	69	75	81	88	94	100																								
17	6	12	18	26	29	35	41	47	53	59	65	71	76	82	88	94	100																							
18	6	11	17	22	28	33	39	44	50	56	61	67	72	78	83	89	94	100																						
19	5	11	16	21	26	32	37	42	47	53	58	63	68	74	79	84	89	95	100																					
20	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100																				
21	5	10	14	19	24	29	33	38	46	48	52	57	62	67	71	76	81	86	90	95	100																			
22	5	9	14	18	23	27	32	36	41	45	50	55	59	64	68	73	77	82	86	91	95	100																		
23	4	9	13	17	22	26	30	35	39	43	48	52	57	61	65	67	74	78	83	87	91	96	100																	
24	4	8	13	17	21	25	29	33	38	42	46	50	54	58	63	67	71	75	79	83	88	92	96	100																
25	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92	96	100															
26	4	8	12	15	19	23	27	31	35	38	42	46	50	54	58	62	65	69	73	77	81	85	88	92	96	100														
27	4	7	11	15	19	22	26	30	33	37	41	44	48	52	56	59	63	67	70	74	78	81	85	89	93	96	100													
28	4	7	11	14	18	21	25	29	32	36	39	43	46	50	54	57	61	64	68	71	75	79	82	86	89	93	96	100												
29	3	7	10	13	17	21	24	28	31	34	38	41	45	48	52	55	59	62	66	69	72	76	79	83	86	90	93	97	100											
30	3	7	10	13	17	20	23	27	30	33	37	40	43	47	50	53	57	60	63	67	70	73	77	80	83	87	90	93	97	100										
31	3	6	10	13	16	19	23	26	29	32	35	39	42	45	48	52	55	58	61	65	68	71	74	77	81	84	87	90	94	97	100									
32	3	6	9	12	16	19	22	25	28	31	34	38	41	44	47	50	53	56	59	63	66	69	72	75	78	81	84	88	91	94	97	100								
33	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	52	55	58	61	64	67	70	73	76	79	82	85	88	91	94	97	100							
34	3	6	9	11	15	18	21	24	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77	79	82	85	88	91	94	97	100						
35	3	6	9	11	14	17	19	23	26	29	31	34	37	40	43	46	49	51	54	57	60	63	66	69	71	74	77	80	83	86	89	91	94	97	100					
36	3	6	8	11	14	17	19	22	25	28	31	33	36	39	42	44	47	50	53	56	58	61	64	67	69	72	75	78	81	83	86	89	92	94	97	100				
37	3	5	8	11	14	16	19	22	24	27	30	32	35	38	41	43	46	49	51	54	57	59	62	65	68	70	73	76	78	81	84	86	89	92	95	97	100			
38	3	5	8	10	13	16	18	21	24	26	29	32	34	37	39	42	45	47	50	53	55	58	61	63	66	68	71	74	76	79	82	84	87	90	92	95	97	100		
39	3	5	8	10	13	15	18	21	23	26	28	31	33	36	38	41	44	46	49	51	54	56	59	62	64	67	69	72	74	77	79	82	85	87	90	92	95	97	100	
40	3	5	8	10	13	15	18	20	23	25	28	30	33	35	38	40	43	45	48	50	53	55	58	60	63	65	68	70	73	75	78	80	83	85	88	90	93	95	98	100

Scoring Examples

E

Third Person Singular Probe

	Correct	Incorrect	Unscorable
Dentist	A dentist cleans your teeth. (He) works. A dentist checks people's teeth. She cleans you teef off. It looks in somebody's mouth.	He fix you. A dentist help our teeth. He pull teeth out. He help your teeth. He grab these.	Dentist doin' here. Dentist teeth. Take teeth out. Teeth
Police Officer	He works in his office. (He) helps people cross the street. A police gets the bad guys. He puts us in jail. He rides bike.	He help that. He catch people. A police say hello. He talk.	Stop. Cop. A bike. He's talking to a little girl.
Firefighter	Firefighter sprays on the fire. He squirts right there. He puts the fire down. Fireman fights people. (A firefighter) saves kids.	He get those fires out. He turn that water on. He fight fires.	They sprayed water. Go away. He fired it. He a fire home. Water blowing it. Fire, fire.
Pilot	She drives a plane. She flies it. He puts the wings up. He goes up in the air. It gets to fly around.	He go fly. A pilot fly. She fly airplane. He stop.	We already ride a plane. Flying on the moon. Airplane. Pilot.
Painter	(She) paints on her house. He gets yellow out. A painter paints. She spreads all the paint.	She paint a house. She wash the house. She color it.	They paint stuff. She painting. She's painting. Paint do house.
Baseball Player	He plays baseball. He hits bat. He hits the ball. He runs fast to base.	He run fast. Baseball player play balls. He throw the ball. He get the ball.	He hit the ball. I have a bat. Player with ball. They throw it. He batting.
Nurse	She gives medicine. She puts on bandaids. She helps you when you are hurt. She puts that on your leg.	A nurse help people. She hurt. She take a bandaid off.	That's a boo-boo. Bandaids. She's fixing the girl. Nurse
Astronaut	He flies. A astronaut flies in a rocket. She goes in rocket ships. She puts her helmet on her head. A astronaut floats.	He get off the spaceship. He go up fast. She fly up high. She come to earth.	This is our airplane. They fly up in the air. He's fly the astronaut. That one.
Dad	He plays softball with his son. He catches a ball. A dad plays with a girl. He throws the ball at the girl.	He throw the ball. He play baseball. He go in the house.	He's playing football. He has that on. Play. Get ball. That's our house.
Dancer	She dances. Her puts her leg up. He spins around. A ballet-er dances.	She go "tada". She dance. It wear a "tu-tu."	The dancer. Foot high. He dancing. They dance. It's a ballerina. She do ballet.

Past Tense Probe

	Regular Correct	Regular Incorrect	Irregular Correct	Irregular Incorrect	Overregularization	Unscorable
painted	He painted the gate. He *colored* it with paint. He *changed* the color.	He paint it. He *change* it.				He is done painting. He paints the fence.
caught	She *grabbed* it.		She caught the ball. She *got* the ball in her glove. She *fell* down.	She catch the ball.	She catched the ball.	Do you know how to catch?
made			She made a birdhouse. He *built* it. He *hung* it on the tree.	She make it. He *build* a house for the birds.	He maked a birdhouse for the birds. He *hanged* it on a tree.	He put the birdhouse on the tree. Looking at the birdhouse.
brushed	He brushed the hair. He *looked* in the mirror. He *combed* his hair.	He brush his hair. He *comb* it.				I can comb my hair. My mommy does my hair.
cleaned	He cleaned the room. She *tidied* up.	She clean the room.				She is going to play now.
kicked	She kicked it high. She *lifted* her leg high.	She kick it.				
wrote	She *wanted* to write the word.		She wrote it.	She write on the board. She *draw* with chalk.	She writed "bell" on the board. She *drawed* on the board.	
climbed	She climbed the ladder.	She climb.	She *went* up the ladder.			She is done.
jumped	He jumped in the puddle. He *splashed* in the water.	He jump. He hop in the water.				He jumps in the puddle.
rode	He *tied* it on the wood. He *stopped* riding.	He tie up the horse.	He rode the horse. He *got* off the horse.	He ride the horse. He get off the horse.	She rided a horse.	
picked	He picked the flowers. He *gathered* them up. He *smelled* the flowers.	He pick the flower.	He *got* them all. He *took* them out.			
dug	She *used* the shovel in the sand.		She dug the hole. She *made* a pile.	She dig in the sand.	She digged a hole. She *maked* a pile.	

	Regular Correct	Regular Incorrect	Irregular Correct	Irregular Incorrect	Overregularization	Unscorable
planted	She planted the flowers. She already planted.	She plant them.				He did some flowers.
ate			He ate them. He *drank* his milk.	He eat them.	He eated a cookie. The boy ated it.	It's all gone. He eats them. I like cookies. I have cookies at home.
blew			She blew out the candle.	She blow out a candle.	She blowed it out. She's gonna eat the cake cause she blowed out the candle.	
tied	She tied her shoe. Her tied her shoe. She *laced* it up.	She tie the shoe.				She's done lacing.
lifted	She lifted the box with both hands. She *picked* it up.	She lift it.				She was lifting. She's carrying the box. She is strong.
gave	She *wrapped* the present.		He gave it to his mom.	He give it to his mom.	He gived the flowers in a box. She gaved it to her mother.	The mother is opening the present. He did give her the present.

Notes regarding scoring examples:

1. This list is not exhaustive but is meant to give you a framework or pattern of responses to use when scoring a child's responses.
2. Italicized verbs indicate a substitution by the child. Note that although children will most frequently use the target stem in their responses, substitutions are acceptable and are scored on the correctness of that particular verb. Also note that substitutions of a different verb type, although they are not optimal and warrant additional prompts during administration, are ultimately also scored as correct or incorrect for that particular verb type.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discrimination function analysis. *Journal of Speech, Language, and Hearing Research*, 41, 1185–1192.

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics*, 15, 507–550.

Brown, R. (1973). A first language. Cambridge, MA: Harvard University Press.

Cleave, P. L., & Rice, M. L. (1997). An examination of the morpheme BE in children with specific language impairment: The role of contractibility and grammatical form class. *Journal of Speech, Language, and Hearing Research*, 40, 480–492.

Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 211–232). Cambridge, MA: MIT Press.

Hall, W. S., Nagy, W. E., & Linn, R. (1984). Spoken words: *Effects of situation and social group on oral word usage and frequency*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Hansson, K., Nettelbladt, U., & Leonard, L. B. (2000). Specific language impairment in Swedish: The status of verb morphology and word order. *Journal of Speech, Language, and Hearing Research*, 43, 848–864.

Haskill, A. M., Tyler, A. A., & Paul, K. (2001, June). Generalization in treatment of the morpheme BE. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., Douglas, L., Taback, N., Lam, I., & Wang, M. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research* 42, 744–760.

Kamhi, A. G., & Koenig, L A. (1985). Metalinguistic awareness in normal and language-disordered children. *Language, Speech, and Hearing Services in Schools*, 16, 199–210.

Leadholm, B., & Miller, J. (1992). *Language sample analysis: The Wisconsin guide*. Milwaukee: Wisconsin Department of Public Instruction.

McDaniel, D., & Cairns, H. S. (1996). Eliciting judgments of grammaticality and reference. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 233–254). Cambridge, MA: MIT Press.

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech, Language, and Hearing Research*, 44, 165–178.

Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research*, 44, 207–223.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research*, 44, 834–847.

Paradis, J., Crago, M., Genesee, F., & Rice, M. L. (2001, June). *French-English Bilingual children with SLI: How Do They compare with Their Monolingual Peers?* Poster presented at the Symposium for Research on Children with Language Disorders, Madison, WI.

Redmond, S. M., & Rice, M. K. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research*, 41, 688–700.

Rice, M. L. (2000). Grammatical symptoms of specific language impairment. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. East Sussex, UK: Psychology Press Ltd.

Rice, M. L., & Bode, J. (1993). Gaps in the verb lexicons of children with specific language impairment. *First Language*, 13, 113–131.

Rice, M. L., Mervis, C., Klein, B. P., & Rice, K. J. (1999, November). *Children with Williams syndrome do not show an EOI stage*. Paper presented at the Boston University Conference on Language Development, Boston.

Rice, K. J., Rice, M. L., & Redmond, S. M. (2000, June). *MLU outcomes for children with and without SLI: Support for MLU as a matching criterion*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Rice, M. L., Spitz, R. V., & O'Brien, M. (1999). Semantic and morphosyntactic language outcomes in biologically at-risk children. *Journal of Neurolinguistics*, 12, 213–234.

Rice, M. L., & Tomblin, B. (1999, June). *Clinical indices of language impairment: Grammatical tense compared to conventional testing*. Poster presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research*, 39, 1239–1257.

Rice, M. L., Wexler, K., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, 38, 850–863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairments. *Journal of Speech, Language, and Hearing Research*, 41, 1412–1431.

Rice, M. L., Wexler, K., Marquis, J., & Hershberger, S. (2000). Acquisition of irregular past tense by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 43, 1126–1145.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 42, 943–961.

Scarborough, H. S. (1990). Index of productive syntax. *Applied Psycholinguistics*, 11, 1–22.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*, 42, 943–962.

Tager-Flusberg, H. & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research*, 42, 1275–1278.

Thornton, R. (1996). Elicited production. In D. McDaniel, C. McKee, & H. C. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 77–102). Cambridge, MA: MIT Press.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, Z., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research* 40, 1245–1260.

Watkins, R. V., Rice, M. L., & Moltz, C. C. (1993). Verb use by language-impaired and normally developing children. *First Language*, 13, 113–131.

Wexler, K. (1992). *Optional infinitives, head movement, and the economy of derivation in child grammar*. (Occasional Paper No. 45). Center for Cognitive Science, Massachusetts Institute of Technology, Cambridge.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (Eds.), *Verb movement*. Cambridge, England: Cambridge University Press.

Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. L. Rice (Ed.), *Toward a genetics of language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Zhang, S., & Tomblin, J. B. (2000). The association of intervention receipt with speech-language profiles and social-demographic variables. *American Journal of Speech-Language Pathology*, 9, 345–357.