
Part I presents an overall rationale for the early identification of strengths and possible developmental weaknesses in school beginners. This is done with the research-based areas of assessment of tests for auditory analysis skills, conceptual skills and language skills. The general purposes of diagnostic assessment and a model of a behavioral assessment funnel for early identification and intervention are presented. The validity and reliability of the tests are discussed.

Part II concerns each test's psychological and educational aim, physical characteristics and development. Administering procedures, research background and score interpretation are presented with implications for educators and parents. Part III is a theoretical and technical discussion of the underlying model of measurement in the series. The appendices include a technical glossary, a table of correlations and reliability coefficients, suggestions for intervention and remediation and a word source list for the word knowledge test. Tables for converting raw scores to ability estimates are given.

Accompanying the handbook are test copies, directions for administration and score keys on:
1. word knowledge
2. comprehension
3. negation
4. verb tense
5. pronouns
6. prepositions
7. figure formation
8. numbers
9. recognition of initial consonant sounds
10. auditory discrimination.

EDRS Price: MF01 Plus Postage. PC Not Available from EDRS.

Descriptors: Cognitive Tests; Educational Diagnosis; Foreign Countries; Learning Readiness; Predictive Measurement; Primary Education; Screening Tests; Student Evaluation

Identifiers: Australia; Australian Council for Educational Research; Test Batteries

Abstract: An overall rationale for the early identification of strengths and possible developmental weaknesses in school beginners is presented in Part I, with the research-based areas of assessment of tests for auditory analysis skills, conceptual skills and language skills. The general purposes of diagnostic assessment and a model of a behavioral assessment funnel for early identification and intervention are presented. The validity and reliability of the tests are discussed. Part II concerns each test's psychological and educational aim, physical characteristics and development. Administering procedures, research background and score interpretation are presented with implications for educators and parents. Part III is a theoretical and technical discussion of the underlying model of measurement in the series. The appendices include a technical glossary, a table of correlations and reliability coefficients, suggestions for intervention and remediation and a word source list for the word knowledge test. Tables for converting raw scores to ability estimates are given. Accompanying the handbook are test copies, directions for administration and score keys on: (1) word knowledge, (2) comprehension, (3) negation, (4) verb tense, (5) pronouns, (6) prepositions, (7) figure formation, (8) numbers, (9) recognition of initial consonant sounds and (10) auditory discrimination.

Reproductions supplied by EDRS are the best that can be made from the original document.
Early Identification and Intervention

A HANDBOOK FOR TEACHERS AND SCHOOL COUNSELORS

Helga A.H. Rowe

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

D. Maguire

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"
## CONTENTS

<table>
<thead>
<tr>
<th>Part</th>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td></td>
<td>ix</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td><strong>Part I</strong></td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>The Content and Structure of this Handbook</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>A New Approach to Early Identification and Intervention</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>The ACER Early School Series</td>
<td>12</td>
</tr>
<tr>
<td><strong>Part II</strong></td>
<td>The Tests</td>
<td>29</td>
</tr>
<tr>
<td>Area A</td>
<td>Tests of Auditory Analysis Skills</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Auditory Discrimination Test</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>Recognition of Initial Consonant Sounds Test</td>
<td>41</td>
</tr>
<tr>
<td>Area B</td>
<td>Tests of Conceptual Skills</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>Number Test</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>Figure Formation Test</td>
<td>58</td>
</tr>
<tr>
<td>Area C</td>
<td>Tests of Language Skills</td>
<td>67</td>
</tr>
<tr>
<td>8</td>
<td>Tests of Syntax</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>Word Knowledge</td>
<td>87</td>
</tr>
<tr>
<td><strong>Part III</strong></td>
<td>Further Technical Information</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>The Measurement Model Adopted for the ACER Early School Series</td>
<td>99</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td></td>
<td>114</td>
</tr>
<tr>
<td><strong>Appendixes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Glossary of Terms</td>
<td>125</td>
</tr>
<tr>
<td>II</td>
<td>Point-biserial Correlations of All Items, and KR-20 Reliability Coefficients for Each Test</td>
<td>133</td>
</tr>
<tr>
<td>III</td>
<td>Suggestions Concerning Intervention and Remediation</td>
<td>135</td>
</tr>
<tr>
<td>IV</td>
<td>Word Knowledge: Source List of Words</td>
<td>138</td>
</tr>
<tr>
<td>V</td>
<td>Conversion of Raw Scores to Rasch Calibrated Ability Estimates for All Tests</td>
<td>142</td>
</tr>
</tbody>
</table>
### Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Multi-skill Multi-sample Matrices for the Auditory Analysis and Conceptual Skills Tests</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>Multi-skill Multi-sample Matrices for the Tests of Syntactic Structures</td>
<td>26</td>
</tr>
<tr>
<td>4.1</td>
<td>Auditory Discrimination Test: Performance of 194 School Beginners</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Auditory Discrimination Test: Item DifficultiesExpressed as Percentage Correct</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Auditory Discrimination Test: Rasch-scaled Items, Difficulty Estimates (Di), and Standard Errors (SEDi) for Three Independent Samples</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Auditory Discrimination Test: Rasch-scaled Ability Estimates (Ai) and Standard Errors (SEAi) for Three Independent Samples</td>
<td>38</td>
</tr>
<tr>
<td>4.5</td>
<td>Auditory Discrimination Test: Correlations between Item Difficulties and between Ability Estimates for Three Independent Samples of School Beginners</td>
<td>40</td>
</tr>
<tr>
<td>5.1</td>
<td>Recognition of Initial Consonant Sounds Test: Performance of 197 School Beginners</td>
<td>44</td>
</tr>
<tr>
<td>5.2</td>
<td>Recognition of Initial Consonant Sounds Test: Item DifficultiesExpressed as Percentage Correct</td>
<td>45</td>
</tr>
<tr>
<td>5.3</td>
<td>Recognition of Initial Consonant Sounds Test: Rasch-scaled Item Difficulty Estimates</td>
<td>46</td>
</tr>
<tr>
<td>5.4</td>
<td>Recognition of Initial Consonant Sounds Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi) for Five Independent Samples</td>
<td>48</td>
</tr>
<tr>
<td>5.5</td>
<td>Recognition of Initial Consonant Sounds Test: Correlations between Ability Estimates Obtained for Five Independent Samples of School Beginners</td>
<td>49</td>
</tr>
<tr>
<td>6.1</td>
<td>Number Test: Performance of 192 School Beginners</td>
<td>54</td>
</tr>
<tr>
<td>6.2</td>
<td>Number Test: Item DifficultiesExpressed as Percentage Correct</td>
<td>55</td>
</tr>
<tr>
<td>6.3</td>
<td>Number Test: Rasch-scaled Item Difficulty Estimates</td>
<td>56</td>
</tr>
<tr>
<td>6.4</td>
<td>Number Test: Rasch Ability Estimates</td>
<td>57</td>
</tr>
<tr>
<td>7.1</td>
<td>Figure Formation Test: Performance of a Representative Sample of 193 School Beginners</td>
<td>62</td>
</tr>
<tr>
<td>7.2</td>
<td>Figure Formation Test: Item DifficultiesExpressed as Percentage Correct</td>
<td>63</td>
</tr>
<tr>
<td>7.3</td>
<td>Figure Formation Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi) for Five Independent Samples</td>
<td>64</td>
</tr>
<tr>
<td>7.4</td>
<td>Figure Formation Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi) for Five Independent Samples</td>
<td>64</td>
</tr>
<tr>
<td>7.5</td>
<td>Figure Formation Test: Correlations for Item Difficulties (Di) and for Ability Estimates (Ai) Obtained between Five Independent Samples</td>
<td>65</td>
</tr>
<tr>
<td>8.1</td>
<td>Prepositions Test: Performance of 217 School Beginners Five Months after the Beginning of the School Year</td>
<td>75</td>
</tr>
</tbody>
</table>
8.2 Verb Tense Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

8.3 Pronouns Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

8.4 Negation Test: Performance of 216 School Beginners Five Months after the Beginning of the School Year

8.5 Comprehension Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

8.6 Tests of Syntax: Order of Difficulty from Easiest to Most Difficult

8.7 Prepositions Test: Item Difficulties Expressed as Percentage Correct

8.8 Verb Tense Test: Item Difficulties Expressed as Percentage Correct

8.9 Pronouns Test: Item Difficulties Expressed as Percentage Correct

8.10 Negation Test: Item Difficulties Expressed as Percentage Correct

8.11 Comprehension Test: Item Difficulties Expressed as Percentage Correct

8.12 Prepositions Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi)

8.13 Verb Tense Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi)

8.14 Pronouns Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi)

8.15 Negation Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi)

8.16 Comprehension Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi)

8.17 Prepositions Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi)

8.18 Verb Tense Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi)

8.19 Pronouns Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi)

8.20 Negation Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi)

8.21 Comprehension Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi)

8.22 Tests of Syntactic Structure: Correlations between Item Difficulties (Di) and between Ability Estimates Obtained in Two Independent Samples

9.1 Word Knowledge Test: Performance of 215 School Beginners after Five Months of Schooling

9.2 Word Knowledge Test: Item Difficulties Obtained from a Sample of 215 School Beginners

9.3 Word Knowledge Test: Rasch Item Difficulties (Di) and Standard Errors (SEDi) of Item Difficulty Estimates
9.4 Word Knowledge Test: Conversion of Raw Score to Rasch-scaled Word Knowledge Score (A1) and Standard Errors (SE$_{A1}$) 95
10.1 Comparison of Rasch-scaled Item Difficulties Calibrated on Different Groups of Children 106
10.2 Number Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 110
10.3 Figure Formation Test: Ratios of Item Difficulty Values for Paired Items for Five Independent Samples 110
10.4 Prepositions Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 111
10.5 Verb Tense Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 111
10.6 Pronouns Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 112
10.7 Negation Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 112
10.8 Comprehension Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples 113

A.1 Point-biserial Correlations and KR 20 Reliability Coefficients for All Tests 133
A.2 Word Knowledge Test: Source List of Words 138
A.3 Conversion of Raw Score to Ability Estimate 142

Figure
3.1 Model for the Standard Error (SE) 22
This Handbook centres around the theoretical basis, research findings, and application of the tests contained in the ACER Early School Series, but its scope and content extend further. It presents a more general theoretical framework within which early identification procedures can be developed, a set of diagnostic tests based on theory, and a means of translating theory-based diagnostic information into behavioural intervention.

The theoretical chapters of this book and the review of the literature pertaining to each area of assessment are expected to be of general interest to professionals and students concerned with early childhood development. The intention has been to present to the reader some of the available knowledge and research findings concerning the identification of important componential skills of early school learning in young children.
ACKNOWLEDGMENTS

The information presented in this Handbook and the tests making up the ACER Early School Series resulted from several years of research and experimentation. Many teachers, counsellors, and children co-operated in item and test trials. They made the final products possible, and I would like to thank them all. I acknowledge the advice and support received from the members of the advisory committee, Miss J. Alsop, Miss M. Calder, Miss H. Eldridge, Miss B. Frazer, Miss M. George, Mr B. Jacka, Mrs J. Mitchell, and Miss M. Roache. I acknowledge with gratitude the contribution made by the late Mr D. Turner in the validation of the Figure Formation and Number Tests. Finally I am indebted to staff at the ACER, among them Mrs Pat Larsen, who worked on the project for a number of years and wrote most of the items for the language tests.
PART I
INTRODUCTION
THE CONTENT AND STRUCTURE OF THIS BOOK

This Handbook was designed to accompany the 10 tests and test-specific Directions for Administration contained in the ACER Early School Series. It contains an overall rationale for the early identification of strengths and possible developmental weaknesses in school beginners. The choice of areas for assessment is research based. The purpose and rationale for each test are discussed in detail, as are the implications for the scores which might be obtained. The implications of performance, discussed separately for each test, include suggestions for remedial activities and recommendations for other aspects of early intervention. The information contained in the Handbook is presented in three parts.

Part I contains the introductory chapters, 1 to 3, which discuss the structure of the information presented in the Handbook and the aims of early identification and intervention, thus providing a framework for the more detailed presentation in later chapters of the purpose and rationale of the type of assessment provided by the ACER Early School Series. The present chapter explains the structure and use of the Handbook.

Chapter 2 presents the broad aims of early school assessment, and describes the nature of the information which teachers, school counsellors, parents, and others concerned with the education of young children can obtain from the tests. The general purposes of diagnostic assessment and screening are discussed, and Cronbach's (1970) 'behavioural assessment funnel' is utilized as a model for early identification and intervention.

Chapter 3 provides general information concerning the nature, purpose, and development of the tests contained in the ACER Early School Series. Questions of validity and reliability are discussed, and the research samples are described.

Part II of this Handbook contains detailed descriptions of the rationale, development, and interpretation of the tests which are contained in the ACER Early School Series. One chapter is devoted to each test with the exception of the tests of syntactic structures, the discussions of which are combined in a single chapter.

The skills assessed by the tests making up the ACER Early School Series can be
INTRODUCTION

regarded as falling into three broad areas. Two tests relate to each of these areas, and the chapters covering each test are presented in the following order:

Chapter 4 Auditory Discrimination
Auditory Analysis Skills
Chapter 5 Recognition of Initial
Consonant Sounds
Number
Chapter 6 Figure Formation
Conceptual Skills
Chapter 7 "Syntactic Structures"

Language Skills
Chapter 8 "Syntactic Structures"
Figure Formation
Chapter 9 Word Knowledge

The presentation of the tests in these three broad areas is a matter of organization and convenience rather than a strict indication of the aims and contents of the tests. All of the tests measure multi-dimensional variables and multi-determined skills and behaviours.

To facilitate the use of this Handbook, the chapters have been structured as consistently as possible. The beginning of each of the three broad areas contains a general statement concerning the rationale and purpose of their inclusion in the ACER Early School Series. Subsequently each test is described under the following section headings within each chapter:

Rationale
Description
Development
Administration and Scoring
Research Results
Interpretation of Scores
Implications.

The Rationale statement defines the aims of the test and justifies the assessment of the particular sets of skills pertaining to the area on the basis of published psychological and educational research and theory. Major research studies are cited. The rationale statement thus provides the test user with an indication of the psychological and educational reasoning underlying the test, and describes the skills the test is intended to measure.

The Description section of each chapter reports on the physical characteristics of the test and includes, where relevant and necessary, a discussion of particular aspects or characteristics of the items, procedure, scoring, etc.

The Development of each test is described fully. Validation studies and, where appropriate, constraints influencing the development of the test are discussed.

The discussion of Administration and Scoring for each test is concerned with important aspects of the procedure of test administration. This section does not include the general directions for administration nor the verbatim instructions for each item, as this information is provided in the Directions for Administration, a separate booklet which is provided for each one of the 10 tests. The information contained in the Administration and Scoring part of each test chapter of this
Handbook provides useful suggestions for the use of the test, and discusses time requirements and materials.

**Research Results** are provided in the form of a descriptive summary of the performances on each test of described samples of children. Measures of central tendency, traditional test statistics, and validity information based on replication studies are provided. These data are presented, firstly, to provide users of the tests with comparative information which might assist them in the interpretation of the results they may have obtained from their own clients. Secondly the research results obtained provide considerable evidence concerning the validity and the reliability of the tests as tools of measurement.

Further assistance is provided to test users in the **Interpretation of Scores** sections. This part of each chapter provides Rasch-scaled item difficulties and standard errors of measurement, as well as Rasch ability estimates for the raw scores for several independent samples. Standard errors of measurement are given for each ability estimate. To some extent this section represents an extension of the Research Results section. In addition, it provides a means for the test user to interpret obtained results within the framework of a Rasch latent trait model of measurement. Further details about this model, its assumptions, and the fit of the present tests to the model are provided in Chapter 10.

The tables provided in the Interpretation of Scores section of each chapter allow the conversion of raw scores into ability estimates which provide ratio-scaled measurement. A more convenient table, however, which provides a means for converting raw scores from all tests on to a common ability scale is provided in Appendix IV, Table A.2.

The last section of each chapter is entitled **Implications**. This section contains further information which will both assist in the interpretation of test results and lead to intervention and remediation procedures. One of the aims of the tests contained in the ACER Early School Series is to provide teachers, school counsellors, parents, and others concerned with the welfare of the individual child with diagnostic information concerning the child's strengths and weaknesses which will lead to operationally defined educational procedures. General and quite specific suggestions are made in this section of each chapter concerning activities that may facilitate the development of specific skills in which individuals may require special practice or other assistance. To supplement the implications sections of Chapters 4 to 9, Appendix III contains a number of general points and specific suggestions which are applicable to intervention and remediation in all areas.

Part III of the Handbook contains Chapter 10 which provides some theoretical and technical information concerning the model of measurement underlying the ACER Early School Series tests. The choice of model on which to base the development of a test has to be determined by the type of use which is to be made of the test results, and on the nature of the variables or the behaviour which are to be assessed. Chapter 10 contains a brief presentation of traditional item analysis theory and an introduction to the Rasch model of measurement and to the major assumptions underlying this model, and reports some results of an examination of
the fit of the Rasch model assumptions to the tests of the ACER Early School Series.

The bibliography provides full information concerning the sources of the references used throughout the Handbook. Appendix I consists of a glossary of major terms used in this Handbook. Short non-technical definitions are provided for concepts which are commonly used in test manuals, and in the discussion of educational and psychological measurement more generally. Appendix II contains a table showing point-biserial correlations and KR 20 reliabilities for all tests. Appendix III contains a series of general hints for intervention and remediation. Appendix IV contains the source list of words, which provided the item pool for the Word Knowledge Test. The validity and usefulness of any word knowledge test varies with the purpose for which it is used. The decision to include the total list of words in this Handbook was made on the recommendation of teachers, who felt that a comprehensive list of words — which is representative of the vocabulary to which five- and six-year-old children in Australia may be exposed at the present time — would provide useful resource material. The rationale for and the development of the source list is discussed in Chapter 9, where the Word Knowledge Test is described. Appendix V consists of a table which allows for the conversion of raw scores to Rasch calibrated ability estimates for all tests in the ACER Early School Series.

The length of this Handbook may encourage some readers to be selective in the chapters they read or the order in which they read them. For these readers the following suggestions are made:

1. All readers should initially read Chapters 1, 3, and the test chapters, 4 to 9, to gain a general perspective of the aims and limitations of the ACER Early School Series.

2. It is essential that potential users of any of the tests in the ACER Early School Series familiarize themselves with the test(s), i.e. understand the rationale, aims, implications, etc. as presented in the relevant chapter before administering the test(s).

3. The information contained in Chapter 2 is of general interest, and is expected to aid the understanding of the purposes of the series considerably.

4. Parts of Chapter 3, Chapter 10 and Appendix I are of a more technical nature, and might be skimmed only by some test users. However, every effort was made to present the information contained in these chapters in as non-technical a manner as possible. It is suggested, therefore, that, with the help of the definitions provided in Appendix I, both Chapters 3 and 10 should be meaningful to most readers.
A NEW APPROACH TO EARLY IDENTIFICATION AND INTERVENTION

Lack of academic achievement, social disadvantage, and an inability to cope in adult life have frequently been related to immaturity or 'lack of readiness' at school entry. Over the past 30 years many research studies have linked low achievement on perceptual and other pre-reading tests with later academic and social problems.

The need for early identification has been acknowledged, especially in the United States, since the early 1960s, but traditional procedures of early identification have been aimed at the prediction of failure in school in general or in relation to various areas of academic requirements. Too often this initially legitimate concern has led to no more than a categorization and labelling of children early in their school careers, and to self-fulfilling prophecies.

Early identification tests of the traditional type have tended to be IQ-type tests, consisting of tasks which correlated highly with success in reading. Children who obtained low marks on such tests were identified as poor risks, and tended to be assigned to less demanding, frequently less reading-orientated activities than those children who had performed well on the tests. Subsequent evaluation of reading achievement tended to show that the children who had previously obtained the low scores on the early identification tests were, in fact, less advanced in reading, thus validating the early identification procedures (e.g. Bond and Tinker, 1973; Book, 1974; De Hirsch and Jansky, 1966; Fry, 1965).

One might suggest that, rather than being impressed by the predictive validity of the above and other tests, one might be surprised if not distressed that little attempt appears to have been made to thwart such prediction. To alert a parent or teacher to the possibility of a poor risk might be justified in situations where a choice exists which might lead to the avoidance of the risk. In the case of most school beginners in Australia, there is no such choice.

Teachers and other educators are generally less concerned with the prediction of failure than they are with its avoidance. They will, therefore, be anxious to discover why an individual might be a poor risk, and what might be done to help the particular individual to become a better risk.

Teachers have become increasingly aware of the need to adapt teaching methods
and materials to the developmental levels and needs of learners. This awareness has led to a search for reliable and valid measures for the assessment of these needs.

Individually administered Piagetian tasks, although found to be valid indicators of developmental levels in some areas (e.g. Bart, 1971; De Vries, 1974; Goldschmid, 1967; Goldschmid and Bentler, 1968) require highly experienced testers and special equipment, and are too time-consuming for practical classroom use.

In the absence of generally accepted procedures, some teachers have administered predictive tests, and then attempted to interpret the results in diagnostic terms. Others have been persuaded to try out poorly conceived assessment procedures (i.e. tests for which no rationale, validity information, etc. has been provided).

The choice of the test by means of which the teacher will conduct an assessment (i.e. a description of certain characteristics of the child) has to be made by giving consideration to both the use which is to be made of the test results, and the nature of the characteristics or the behaviour which are to be assessed. To the reader of this Handbook, this statement may appear to be a rather obvious one. It seems to have been less obvious to many developers of early identification tests.

New purposes or aims in assessment might require new approaches to the development of tests. Certainly a distinct aim to assess a specific set of variables must be accompanied by a correspondingly distinct conceptualization of the test rationale and possibly a specifically appropriate methodology. In other words, new ways of conceiving the characteristics of learners require altered modes of measuring them, and may require modifications in the procedures and criteria used to establish the quality of these measuring devices as well.

If it is the aim to identify particular strengths and weaknesses of individual children as well as the characteristics of the class as a whole so that the most appropriate teaching procedures can be developed, the teacher requires tests, the performance on which can be translated directly into specific learning requirements. Such tests allow for the interpretation of test performance in terms of specific skills or other modifiable behaviours.

Tests which lead to a description of performance in behavoural terms have much in common with traditional measurement of achievement in the classroom. However, while all tests of achievement assess behaviour, not all allow for the expression of test results in terms of behaviours. The major difference lies in the teacher's purposes for using the test, and in the manner in which the test results are used. The methods of assessment, even some of the item content, may appear to be of the same general type; it is the tailoring of assessment to the specific educational problem which makes the difference.

Phases and Purposes of Assessment
To place the purpose of developmental and educational assessment — specifically the early identification of children who may be developmentally or otherwise disadvantaged — into some perspective, it might be useful to think of the
A NEW APPROACH

A variety of models could be conceptualized but, for the purposes of the present discussion, an arbitrary choice has been made to view the assessment needs of school beginners as involving five phases. These phases are not discrete; they are expected to overlap. In addition to this, recycling to earlier phases is often useful in later phases when new information comes to light. As a model, these phases might be regarded as representative of what Cronbach (1970) called a 'behavioural assessment funnel' because a maximally broad band of assessment narrows rapidly to focus on those specific areas identified during the broad phase, and then maintains a relatively narrow focus through the phases of intervention and continuous monitoring.

The five phases might be perceived as follows:

Screening and assessment of general developmental variables

In this initial phase a general assessment of the developmental and social characteristics is made, and possible areas of difficulty are noted. Developmental variables are considered and decisions are made as to the nature of additional, more precise assessment procedures. The methods used in this phase are generally quite broad, and often of low fidelity (Cronbach, 1970), and might include teachers' first impressions, classroom observation, parent interview data, the ACER Checklists for School Beginners, or other checklists, and similar wide-ranging assessment devices.

Frequently qualitative rather than quantitative data are obtained at this point. The assessment literature tends to ignore this and the following phase, or to treat them quite briefly. In fact little objective evidence appears to be available on the processes involved in determining the targets for early intervention (Hawkins, 1975). The literature tends to focus on how to estimate the severity of a problem, but provides little assistance in determining how the problem is to be selected as crucial. In many instances, for example, two different infant teachers observing the same child would not necessarily select the same problem area for further investigation and intervention.

Definition of the problem

This phase includes two major functions, measurement and hypothesis formation. It is likely that the teacher or counsellor shifts back and forth between these functions several times during this phase. Measurement may be aimed at classifying the problem for administrative records. It may be oriented towards determining what general type of intervention may be required (e.g. skill training, change in class, or both), or towards some other general decision making. An example of measurement in this sense would be the administration of a standardized achievement test which can be used to compare relative performance among peers, or even reading group placement for the child, but which does not lead to the identification of which specific skills the child might need to be taught next.
Hypotheses may be generated as to which behavioural or skill strengths and weaknesses might play a key role in the observed major problem and difficulties; for example, whether observed difficulties in adapting to school might stem from learning, motivation, or social problems (or all of these). This phase includes the assessment of the strengths of the child, and of the educational resources available to him or her.

**Pinpointing and design of intervention**

The third assessment phase requires the teacher to select specific behaviours or environmental variables as targets for change. A narrow band, high fidelity (Cronbach, 1970) assessment procedure is appropriate here. This phase may, therefore, be represented as the narrowest point in the cup of the assessment funnel. The pinpointing aspect of the assessment need not be formal. In fact, the best pinpointing often turns out to be informal.

In this phase, assessment must go beyond formal achievement tests which might provide age and/or grade equivalent norms, to more specific educationally diagnostic assessment of the particular component skills involved in, for example, reading and number work. In the assessment of school beginners or developmentally and otherwise handicapped individuals, the tests contained in the ACER Early School Series are expected to be particularly useful during this phase.

Phase 3 involves the making of tentative plans as to possible early intervention procedures, and includes the evaluation of their feasibility, the revision of plans, and possible further assessment of feasibility.

At this stage the teacher predicts the possible effects of various methods of intervention and programming. Examples of such prediction would include consideration of the child’s motivation for remediation or behaviour change, the relevance, availability, and relative promise of a number of reinforcers, and the child’s strengths in particular skill areas which might be employed. Ideally this phase would not conclude before the teacher has established some sort of baseline pre-intervention profile.

The latter constitutes an essential prerequisite for the following two phases, and for other attempts to evaluate remedial or intervention procedures.

**Monitoring of progress**

During this phase the teacher implements the intervention procedures, while continuing to use the assessment method selected to obtain the baseline data. A relatively continuous measurement of the child’s or group’s progress is required and the procedure must be as economical, in terms of time requirements, as possible. It is essential that the skills, behaviours, and environmental variables measured are relevant to the specific objectives of the intervention procedure, and that they are sensitive to the intervention effects. They should, consequently, be criterion referenced and of high fidelity. The precise definition and measurement of relatively identifiable skills and components of more complex skills provide an excellent means for obtaining data which are sensitive to intervention procedures. These types of measures are provided by the tests of the ACER Early School Series.
Because the same measures are usually used repeatedly over time in this monitoring phase, and because such measures tend to cover a relatively narrow range of skills, behaviours, or environmental variables, this phase of the assessment process may be represented by the neck of the behavioural assessment funnel. A number of specific hints concerning intervention and remediation procedures are presented in Appendix III.

Follow-up

Responsible assessment and intervention procedures usually involve some kind of follow-up after the termination of intervention. This justifies the length of the neck of the assessment funnel. The measures used will appropriately be somewhat narrow in focus still, similar to those used during intervention. In fact, the follow-up measures will usually include those used in the previous phase.

The tests contained in the ACER Early School Series provide useful measures for use during phases 3, 4, and 5 of the above suggested model of assessment and intervention.

All 10 tests are based on a criterion referenced model of measurement, and thus provide teachers with meaningful guidance in the development of individualized teaching procedures, and in determining individual progress. Performance on the tests can be translated directly into specific learning requirements for each child. At a time when a strong focus in the delivery of educational services is on accountability, the ACER Early School Series tests can facilitate the demonstration of program effectiveness.

More specifically the benefits of this set of tests include the following:

(a) for each child, a record of a profile of behaviourally defined and teachable skills which have been accomplished at the time of testing and skills not mastered at that time;

(b) providing the teacher with concrete and operationally defined skills which need to be taught in each of ten important areas of development;

(c) allowing the teacher to pinpoint the educational needs of individual children in these areas;

(d) a means for assessing and recording progress for each child, for groups of children, and for the class as a whole;

(e) a means of both formative and summative evaluation;

(f) the possibility that repeated use of the tests would provide a means of ongoing evaluation and therefore facilitate the identification of appropriate individualized learning objectives.
THE ACER EARLY SCHOOL SERIES

Nature and Purpose
The tests contained in the ACER Early School Series were designed to provide a series of estimates of a child's cognitive development and maturity.

Apart from their application for the identification of particular strengths and needs of individual school beginners, the tests are expected to have wider utility as counselling and guidance tools in the lower grades of primary school and with certain groups of handicapped older persons. The tests provide an appropriate means for the assessment of children who cannot read and for remedial students, as reading is not required for the ACER Early School Series.

Trial testing showed that most children were interested in the tests and appeared to enjoy the items. Therefore the tests contained in the ACER Early School Series may facilitate the establishment of rapport with children in clinical and small group situations. No specialized training or testing experience is required on the part of the teacher or counsellor who administers the tests. The directions for administration are explicit and scoring is objective and requires little time. All tests are completely untimed and are thus power rather than speed tests. Despite this, most of the tests take no more than 15 to 20 minutes. Longer tests are designed to be administered in sections over a period of time. Retesting to monitor development and progress is highly recommended.

Verbalization on the part of the testee is not required. The child is asked to place a cross on one of a number of alternative pictures which corresponds to the stimulus presented by the tester. The fact that the tests do not require verbal responses makes it possible to administer them to small groups, thus providing a reasonably economical means of assessment as regards time.

Another advantage of the lack of a verbal response requirement is that this makes the tests a suitable means for the assessment of speech impaired children (e.g. stutterers, children suffering from delayed speech, forms of expressive aphasia, and certain autistic children) and very withdrawn children.

Although the tests are suitable for administration to small groups of children, individual testing will always give more valid results. It is recommended that
physically and/or mentally handicapped individuals always be tested in one-to-one situations.

Motor-impaired children can be asked to point to the appropriate response picture. In more severe cases of perceptual impairment, the tester can, after presenting the stimulus, point to each response alternative in turn, and elicit a 'yes' or 'no' (nod, etc.) reaction from the testee. The ACER Early School Series tests can thus be administered to any individual of reasonable hearing and vision, who understands the stimulus words and who can communicate a yes/no response in some form.

Culture Fairness

School beginners, even more so than older primary school children, have not yet fully internalized the customs, habits, and ways of the culture which surrounds them. This makes it meaningless to consider the culture fairness of the test items.

Another more valid reason for the disregard of culture fairness lies in the intended purpose of the series. As stated previously, the tests contained in the ACER Early School Series are intended to provide the teacher or parent with an indication of the child's strengths and weaknesses in a number of important learning areas so that intervention can take place before the child becomes disadvantaged.

Activities in the classroom of which the child is a part strongly reflect the culture of a school in Australia. A valid assessment of possible cultural, social, developmental (including language) disadvantage of an individual can only be derived from the use of procedures which are culturally relevant to the situation in which the individual has to function.

Major Constraint

A major constraint imposed on the design of the ACER Early School Series was the demand that the tests be of the paper and pencil type, and that all tests in the series could be administered to small groups of children. This led to the exclusion of a number of relevant areas of assessment, and restricted the assessment to areas of receptive functioning.

Description

All the tests are criterion-referenced, each having its own direct implication for teaching the skills which may be found to be weak or lacking in a particular individual.

The entire set of tests was designed to be administered by teachers, counsellors, kindergarten teachers, and in some cases parents, since they are most likely to benefit from the diagnostic information in their endeavour to provide experiences and teaching which are tailored to the specific needs of the individual child.

Development

The initial impetus for the selection of test areas and items was received from the educational and psychological research literature. Extensive pilot testing of
individuals and small groups of school beginners, kindergarten, first grade, and some handicapped children was followed by the revision of a number of initial items and sets of items, and finally resulted in the inclusion of the following tests:

1. Auditory Discrimination Test
2. Recognition of Initial Consonant Sounds Test
3. Number Test
4. Figure Formation Test
5. Prepositions Test
6. Verb Tense Test
7. Pronouns Test
8. Negation Test
9. Comprehension Test
10. Word Knowledge Test

Research Samples:
The research results presented for each test were based on the performances of a number of different samples. The auditory analysis and conceptual skills tests were administered to a representative sample, and to a number of unselected samples. The scores of the latter were provided by teachers and guidance officers who had shown a special interest in the project.

The representative sample consisted of 50 Victorian schools and included government, Catholic, and private schools. This sample was drawn by the survey section of the ACER. Four of the 50 schools were unwilling to participate and were replaced by schools from a matched list. The school principals who had consented to participate in the study were asked to submit complete class lists for the preparatory grades of their schools, and were requested to appoint a co-ordinator who would be responsible for the administration of the tests.

Six children were selected randomly from the class lists submitted by each school. A notice was prepared and forwarded to each participating school, listing the names of four children who were to be tested, and two supplementary names to be used as substitutes in case of absences. The tests were administered in all schools during the first week of August. Forty-nine schools returned the test booklets as requested. The total representative sample thus consisted of 196 school beginners, all of whom were tested on the Auditory Discrimination Test, the Recognition of Initial Consonant Sounds Test, the Number Test and the Figure Formation Test.

The previously mentioned additional, unselected samples differed in number and size for the four tests. Some children in these samples had completed several tests. As most children appeared not to have attempted all four tests, and as the combinations of tests which had been administered together varied, the test performances of these samples were investigated for each test separately. The sizes of the samples of unselected school beginners were substantial, thus providing a good opportunity to validate the observations made on the basis of the representative samples. For the Auditory Discrimination Test, data of an additional 599 school beginners were investigated. The performance of an additional 600
children was examined for the Recognition of Initial Consonant Sounds Test. The scores for an additional 612 and 600 school beginners were analysed for the Number Test and Figure Formation Test respectively.

Because of staffing circumstances at the ACER, it was not possible to administer the language skills tests to the above described representative sample. The sample for the language tests was drawn from five Victorian schools, representing extremes in terms of socio-economic status, and to children in a country school. Two of the schools were situated in the inner suburban area of Melbourne, which has a high incidence of children from non-English language backgrounds. Languages spoken in the families of the children in the sample included Italian, Greek, Macedonian, Portuguese, Serbo-Croatian, Turkish, and Vietnamese. One of the schools was situated in a high socio-economic suburb of Melbourne, one in a high socio-economic area in a provincial city not far from Melbourne, and one in a country area. The total sample consisted of 220 school beginners. Testing took place four months after the beginning of the school year. The performances of children from English and non-English language backgrounds were examined separately as well as in combination.

Application for Diagnosis and Screening

The tests contained in the ACER Early School Series can be used for both purposes of educational diagnosis and screening. Both the terms 'diagnosis' and 'screening' are more frequently used in medical settings. This does not detract from their usefulness in the present context as long as both concepts are suitably defined in educational terms.

The definition of the term 'diagnosis' provided in the Shorter Oxford English Dictionary is 'to identify by careful observation'. This general definition is equally valid for medical, educational, social, and other purposes, in which distinctive characteristics need to be identified. In its application to medicine, the same dictionary defines diagnosis as 'the determination of the nature of a diseased condition' or the 'identification of a disease by investigation of its symptoms and history'. In other words, medical diagnosis is concerned with the etiology and the classification of symptoms. In medical practice, knowledge of the etiology of the disorder is usually essential for both treatment and prognosis.

Unfortunately many research workers and practitioners have attempted to transplant the medical model of the concept of diagnosis into the field of the behavioural sciences, particularly into the areas of learning disabilities and disorders of behaviour, without redefinition. Academic and other school problems were regarded as if they were diseases for which an etiology had to be found which would, ultimately, lead to a specific form of treatment. From this orientation resulted the emphasis on the importance of diagnostic measures, and the belief that specific causes within the individual would be found for specific and more generalized learning difficulties, and for psychological problems.

The early workers concerned with the identification of specific learning disabilities (e.g. Kirk, Clements, the Orton Society, and others publishing in the 1950s) tended to regard dyslexia basically like any physical illness. The only
difference in this view between a condition like dyslexia and a bodily disease was that the former, affecting learning ability, manifests itself by means of functioning in the school situation, whereas the latter, affecting physical organs such as the heart, lungs, skin, etc., manifests itself by means of symptoms relating to those parts of the body.

A considerable and, we are told, increasing number of children fail to benefit optimally from normal classroom instruction. Among them we find both the gifted and those with learning problems. A traditional reaction to the child with the latter problem, once the possibility of laziness has been discounted, tends to be to investigate what is wrong with him or her. There may be nothing wrong with the child. His or her development and performance may be perfectly appropriate in relation to the experiences provided and the constraints imposed by the individual’s environment up to the point of assessment.

Any adult or child may develop learning problems depending on the situation or environment in which he finds himself. On the other hand, a handicapped person’s ability and performance may improve if the environment is structured in such a way as to meet the individual’s specific needs.

The individual with a visual deficit is provided with glasses, the deaf child with a hearing aid. Good provisions tend to be made for a variety of physical handicaps. In the majority of cases our intervention is based on valid and reliable diagnosis: our intervention procedures meet the diagnosed needs of the individual. Is our record equally good in diagnosis and intervention in other educational areas? Teaching, remedial teaching in particular, should be based on the knowledge of individual needs and differences which are operationally defined.

While it is possible that the etiology bound medical model of diagnosis may well be appropriate in the investigation of some learning disorders (e.g. in relation to neurological disease, in certain psychoses like schizophrenia, and in some types of autism; where research appears to be establishing the important role of genetically determined biochemical factors) it seems equally certain that an educational model will prove to be more effective in dealing with the large group of problems subsumed under the heading, ‘learning disabilities’.

One of the most important and detrimental consequences of the transposition of the medical model to education and educational counselling is the attempt to derive diagnostic labels for the classification of individuals. Major objections which can be brought against this practice include its lack of validity and reliability and its disregard of environmental factors. In addition to this, even where consideration was given to the child’s environment and if such labelling were both valid and reliable, it cannot be expected to serve a useful purpose in relation to individuals with learning problems, or in relation to the early identification of school beginners who might be at risk for such problems.

As noted previously, a medical diagnosis is useful if it provides information concerning the etiology of the disorder, or if it has implications for treatment or prognosis. If, however, a diagnosis has no implications for treatment or prognosis, then what is its purpose beyond providing a descriptive label for the individual?

From an educational point of view, it seems quite futile to label an individual,
since labels have no etiological or treatment implications. The practice is of dubious value even in relation to research. In fact, it might serve to draw attention away from more important and useful educational activities.

It might be a wasteful but revealing exercise to add up the number of school hours which have been spent in the past 15 years in Australian schools in attempts to arrive at differential diagnoses in relation to learning difficulties and other educational problems. Mischel (1968) and others have documented impressive findings in relation to this activity in the United States of America. Apart from the fact that many of the tests used identified only those children who had previously been identified by less costly means, in the majority of cases no important consequences were found to have resulted from the diagnoses. Why then waste precious school time to arrive at a diagnosis?

Diagnostic testing is justified, however, if it is based on an educational model of diagnosis. The educational model of diagnostic assessment is equally applicable to the identification of the needs of both the gifted and of those with special academic difficulties. 'Educational diagnosis' can be defined as the attempts made to identify immediate and future specific educational needs of the individual, and as a means to provide a basis for the selection of teaching procedures.

The educational model of diagnosis requires a diagnostic test to indicate specific areas of strength and/or weakness in the child's achievement of a given, operationally defined set of skills and to suggest, as a result of this, appropriate means of intervention.

The tests contained in the ACER Early School Series are diagnostic in the sense of an educational model of diagnosis. Each test is aimed to assess educationally definable and modifiable skills in a specific curriculum-related area.

As noted previously, the tests are criterion-referenced. This means that the set of items constituting each test 'can be conceived as a sample from a domain of tasks covering a well defined ... competency' (Van der Linden, 1980, p. 469).

Since Glaser's (1963) seminal paper which provided the first distinction between norm-referenced and criterion-referenced strategies of measurement, many statements defining criterion-referenced testing have appeared in the literature (Gray, 1978; Gronlund, 1973; Hambleton, 1981; Popham and Husek, 1969; Nitko, 1980). The most widely accepted definition is probably reflected in that used by Hambleton which he basically attributes to Popham:

A criterion-referenced test is constructed to assess the performance levels of examinees in relation to a set of well-defined objectives (or competencies). (Hambleton, 1980, p. 421)

Each of the tests in the ACER Early School Series is designed to measure one major objective. The testee's performance on each test is reported separately and provides an indication of his or her achievement on each of the objectives.

Criterion cut-off scores or any other scores which differentiate one group of children from another, or categorize and label children, are not necessary. The purpose of criterion-referenced tests in educational diagnosis is to focus on the identification and educational description of individual differences. The information yielded by diagnostic testing is expected to increase the effectiveness
of educational intervention because it reveals the child’s strengths and weaknesses in relation to relevant academic skills and sub-skills.

The educational target for all school beginners is full mastery of all tasks of the types assessed in the tests of the ACER Early School Series. In the initial stages of schooling for children who may have particular handicaps, or when a screening instrument is required, the teacher may wish to compare an individual child’s or a group’s performance with that of peers. Tables 1 and 2 of each of the test chapters in this Handbook make it possible for the teacher or counsellor to compare the performance of individuals on each test with the performance of a representative sample of school beginners and with the scores obtained by other groups of children. In this way the tests contained in the ACER Early School Series can be used to obtain both ipsative and comparative information on the individual. It should be noted, however, that the performance data provided in the Handbook do not represent norms: Criterion-referenced tests have no norms; instead they signify skills and behaviours which all children should achieve at some stage. The final norm is thus expected to be complete mastery for all children. The research findings presented in the Handbook are intended solely to provide test users with guidelines and encouragement.

For some purposes of assessment, a combination of the criterion-referenced and normative approaches might seem appropriate. Some test constructors have normed criterion-referenced items to provide educators with guidelines as to the age or stage at which children typically tend to perform a range of skills. This approach has been used successfully in a number of developmental scales, including the Bayley Scales, Cattell Infant Scale, Denver Developmental Screening Test, the Gesell Age Norms, and the Vineland Social Maturity Scale. These tests provide behaviourally described developmental ‘milestones’. However, although some criterion-oriented descriptions are provided, the aim of these types of instruments is basically a norm-referenced one. Items may be criterion linked, but they tend to be widely spaced and not necessarily representative of a defined objective or criterion. The tests are designed to compare the performance of one child with that of peers. These tests may provide information concerning the need for educational intervention. They are not designed to provide the specific knowledge required for the design of individualized teaching procedures and other need-based intervention.

The teacher requires tests which provide reference points concerning the child’s performance on specific skills, and some information concerning a sequential order of developmental skills, close enough in increasing level of difficulty so as to be useful in the development of appropriate instruction.

The usefulness of criterion-referenced tests for educators would certainly be enhanced if the items in each area could be presented in a sequential order of difficulty or developmental level which is generalizable across samples. The Rasch calibration of the ACER Early School Series tests constitutes a considerable movement in this direction. However, the adoption of this type of measurement model should not mislead users to assume that all developmental archetypes of skills are invariant, and that particular strengths or weaknesses of the
child and environmental factors such as home background, pre-school experience, quality of teaching, etc. would not affect the difficulty of specific items for individual children. For example, the above described hybrid model of normed criterion-referenced measurement applied to the Word Knowledge Test of the ACER Early School Series would assume that the difficulty level of individual words would be much the same for all children of similar age or grade, and mainly related to maturation. The multicultural community reflected in most-classrooms in Australia shows the lack of practical validity and utility of such a hybrid model of assessment.

In the absence of generalizable analyses of skills into hierarchically sequenced sub-skills, which would provide a ready-made teaching program, the aim of the ACER-Early-School-Series has been to attempt to select items in each skill area in such a way that they are highly representative of teachable sub-skills. This provides the teacher with performance data for each child on a representative sample of sub-skills in each area, and allows the planning of teaching procedures on a broad front. Instead of assessing level or progress for each child by attempting to identify the point on the borderline at which the child appears to begin to have difficulties (as would be possible on the basis of a set of hierarchically sequenced items according to difficulty level), the items are regarded as a sample representative of the population of possible items relating to each specific skill, and the child's performance on each set of skill-related items will provide the teacher with information concerning the strength of the skill and the types of sub-skills in need of further training. In other words, the child's performance on each test focuses the teacher's attention on specific appropriate sub-skills which particular children need to be taught.

Reliability

The consistency with which a test measures is referred to as its reliability.

Reliability is as important for criterion-referenced tests as it is for conventional norm-referenced tests, and it can be determined in much the same way. The notion that "classical" test theory and "traditional" methods of test analysis are inappropriate for criterion-referenced tests is based on a misconception, namely, that scores on a criterion-referenced test show no variability because all who take the test make perfect scores answering all the questions correctly. While this is a theoretical possibility, it almost never actually happens. (Ebel, 1979: p.281)

Obviously both criterion-referenced and norm-referenced tests should be reliable, as should any observations which are subsequently reported.

Indices of equivalence, stability, and internal consistency tend to be considered in the assessment of the reliability of a normed test. Equivalent reliability is established for tests which have parallel forms. Stability refers to retest reliability, and refers to the consistency of measurements yielded by two administrations of the same test to the same group of testees, usually with the intervention of a short period of time. Reliability in terms of internal consistency refers not to the consistency of the test results as a means of assessment, but to characteristics of the items making up the test. Internal consistency coefficients show to what degree the items making up the test are consistent among themselves.
INTRODUCTION

All three reliability indices and combinations of them can be computed for both normed and criterion-referenced tests. However, for criterion-referenced tests, certain constraints operate and some reliability indices cannot be conceptualized in the same way as they are in norm-referenced tests. The reader interested in pursuing these matters is referred to the research literature (e.g., Ebel, 1979; Popham, 1979). Sufiice it here to note that both reliability and validity coefficients for tests are calculated on the basis of the scores of particular groups of individuals who have taken the test. The characteristics of the group affect the values of both coefficients. The heterogeneity of the performance in the group is of particular importance. All other things being equal, the larger the variance in a set of obtained scores, the greater the possibility of a correlation between those scores and another set of scores. Reliability coefficients obtained in homogeneous groups are expected to be lower than those obtained for samples of individuals whose scores on the test are more heterogeneous. Many criterion-referenced measures yield rather homogeneous performance data, particularly in cases where it is possible to define the objective of the test in a quite narrow way, and where the group includes individuals with and without the skill tested. This restriction-of-range problem distorts the meaning of any estimates which are derived by correlational statistics.

'The nature of the decision alternatives will influence the type of reliability procedure we employ' (Popham, 1979, p.146). The most frequently reported reliability estimate is the index of internal consistency. This reliability coefficient reflects the degree to which items within one test are intercorrelated. Although this examination of response homogeneity might be regarded as somewhat redundant operation in relation to a criterion-referenced test, the items of which measure a well-defined reasonably narrow objective, this type of reliability is reported.

The Kuder-Richardson formula 20, usually symbolized as KR 20 (Kuder and Richardson, 1933) was used to estimate the internal consistency of each test. This formula requires that the test be given once only, and calculates the mean of all possible split-halves coefficients (Cronbach, 1951) for the tested group of individual as follows:

\[
KR\ 20 = r_i = \left( \frac{n}{n-1} \right) \frac{s^2_{ti} - \Sigma pq}{s^2_t}
\]

where

- \( r_i \) is the reliability coefficient for the whole test,
- \( n \) is the number or items in the test
- \( s^2_t \) is the variance of total scores in the test
- \( \Sigma pq \) is calculated by summing the products of the proportion of persons who pass \( (p) \) and the proportion of persons who do not pass \( (q) \), which is computed for each item across items.

As can be seen in Table A.1, the reliability coefficients are moderately high. The actual value of the reliability coefficient is a function of the intercorrelations between items and the length of the test. More specifically, as the average inter-item correlation increases and as the number of items increases, the value of KR 20 increases.

The reliability of the individual tests could have been improved by including
additional highly correlated items. This was not done because the gain for the teacher resulting from the relatively weak impact of increased reliability would seem negligible as compared with the additional time and resources required for the administration, scoring and information of longer tests.

Because measures of internal consistency do not refer to the reliability of decisions resulting from a test, but are restricted to a description of the items themselves, Popham (1979) suggests:

For criterion-referenced testing, we might better conceive of internal consistency estimates as a vehicle for verifying the derivative homogeneity of a set of test items. Conceptually internal consistency approaches are not particularly helpful when thinking of a criterion-referenced test's consistency of measurement. (p.155)

Indices of equivalence are not relevant in the present context, as no parallel forms of the tests are available at present. The stability of both item difficulties and ability estimates are high as shown in the Rasch analyses of the performance data obtained from a number of independent samples of school beginners.

**Standard Errors of Measurement**

Tests are not perfectly reliable. On a retest the scores of most children would probably differ somewhat from their original scores. These changes can result from a variety of variables mainly connected with the testing situation, and are not expected to be criterion related.

The Standard Error of Measurement (SEM) is a statistic which provides the test user with an indication of the amount of change in score which might be expected on a retest. In a retest situation, without intervening learning period, the mean change in score is assumed to be zero, and the changes which do occur are assumed to be distributed according to a bell-shaped curve around zero. In other words, it is assumed that as many children obtain an increased as a decreased score in the second testing. The Standard Error (SE) represents the standard deviation of the changes in score which occur in such a test-retest study.

Figure 3.1 provides an illustration of the general model for the SE, together with its specific application to the tests contained in the ACER Early School Series.

The SE provides a number of benefits to the test user. Most importantly, it provides an indication of the limitations of the reliability of the items or test. It thus reminds test users to be less rigid in their interpretation of test scores. Not only does it show that an obtained test score of 30 might lie somewhere between 29 and 31, but it makes possible a more valid interpretation of performance within the same test battery. In fact, the Rasch calibrated item difficulties and ability estimates are accompanied by an indication of the SE for each ability estimate.

The conventional use made of the SE is to treat each score X as X±1 SE. In this way each test score is interpreted as a range or band, rather than as a fixed point. Where the ranges of scores overlap, the differences between scores are disregarded (i.e. not interpreted as differences in ability). For example, for any item with a SE of 2.5 points, a child's two scores (or the scores of two children) have to be at least five points apart before a difference is interpreted.

**Validity**

It is generally accepted that test validity is not an inherent or fixed attribute of the
test itself, but rather that the validity of a test depends on the use made of it (American Psychological Association, 1974; Cronbach, 1971; Shepard, 1980).

Twenty years ago, Ebel pointed to the fact that the validity of measurement seems not to be a problem in the physical sciences:

> Is it reasonable to judge the quality of a barometer solely or even mainly by the accuracy of the weather forecasts which are made with its help? (Ebel, 1961, p.644)

The requirements of physical measurement are operational definitions of what is to be measured and accuracy of measurement (i.e. reliability). Validity appears to be taken for granted. This is not advocated for developmental and educational measurement at this stage. It is suggested, however, that the ultimate assessment of the validity of a test must be based on whether the results of the individual's performance on the test accomplish the purpose for which the test was intended.

The term 'validity' generally refers to the extent to which a test measures what it purports to measure. Strictly speaking, 'one validates not a test, but an interpretation of data arising from a specified procedure' (Cronbach, 1971, p.447).

### Figure 3.1 Model for the Standard Error (SE)

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>-0.9</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>-1.2</td>
</tr>
<tr>
<td>Number</td>
<td>-0.3</td>
</tr>
<tr>
<td>Figure</td>
<td>-0.3</td>
</tr>
<tr>
<td>Formation</td>
<td>-0.3</td>
</tr>
<tr>
<td>Prepositions</td>
<td>-0.3</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>-0.3</td>
</tr>
<tr>
<td>Pronouns</td>
<td>-0.6</td>
</tr>
<tr>
<td>Negation</td>
<td>-0.3</td>
</tr>
<tr>
<td>Comprehension</td>
<td>-0.3</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>-3.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>-0.6</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>-0.8</td>
</tr>
<tr>
<td>Number</td>
<td>-0.2</td>
</tr>
<tr>
<td>Figure</td>
<td>-0.2</td>
</tr>
<tr>
<td>Formation</td>
<td>-0.2</td>
</tr>
<tr>
<td>Prepositions</td>
<td>-0.2</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>-0.2</td>
</tr>
<tr>
<td>Pronouns</td>
<td>-0.4</td>
</tr>
<tr>
<td>Negation</td>
<td>-0.2</td>
</tr>
<tr>
<td>Comprehension</td>
<td>-0.2</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>-2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>-0.3</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>-0.4</td>
</tr>
<tr>
<td>Number</td>
<td>0</td>
</tr>
<tr>
<td>Figure</td>
<td>0</td>
</tr>
<tr>
<td>Formation</td>
<td>0</td>
</tr>
<tr>
<td>Prepositions</td>
<td>0</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>0</td>
</tr>
<tr>
<td>Pronouns</td>
<td>0</td>
</tr>
<tr>
<td>Negation</td>
<td>0</td>
</tr>
<tr>
<td>Comprehension</td>
<td>0</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>0</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>+0.4</td>
</tr>
<tr>
<td>Figure</td>
<td>+0.1</td>
</tr>
<tr>
<td>Formation</td>
<td>+0.1</td>
</tr>
<tr>
<td>Prepositions</td>
<td>+0.1</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>+0.1</td>
</tr>
<tr>
<td>Pronouns</td>
<td>+0.2</td>
</tr>
<tr>
<td>Negation</td>
<td>+0.1</td>
</tr>
<tr>
<td>Comprehension</td>
<td>+0.1</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>+1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>+0.3</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>+0.8</td>
</tr>
<tr>
<td>Number</td>
<td>+0.2</td>
</tr>
<tr>
<td>Figure</td>
<td>+0.2</td>
</tr>
<tr>
<td>Formation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Prepositions</td>
<td>+0.2</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>+0.2</td>
</tr>
<tr>
<td>Pronouns</td>
<td>+0.4</td>
</tr>
<tr>
<td>Negation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Comprehension</td>
<td>+0.2</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>+2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>+0.6</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>+1.2</td>
</tr>
<tr>
<td>Number</td>
<td>+0.9</td>
</tr>
<tr>
<td>Figure</td>
<td>0</td>
</tr>
<tr>
<td>Formation</td>
<td>0</td>
</tr>
<tr>
<td>Prepositions</td>
<td>0</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>0</td>
</tr>
<tr>
<td>Pronouns</td>
<td>0</td>
</tr>
<tr>
<td>Negation</td>
<td>0</td>
</tr>
<tr>
<td>Comprehension</td>
<td>0</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>0</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>+0.8</td>
</tr>
<tr>
<td>Figure</td>
<td>+0.6</td>
</tr>
<tr>
<td>Formation</td>
<td>+0.4</td>
</tr>
<tr>
<td>Prepositions</td>
<td>+0.2</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>+0.2</td>
</tr>
<tr>
<td>Pronouns</td>
<td>+0.4</td>
</tr>
<tr>
<td>Negation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Comprehension</td>
<td>+0.2</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>+3.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>0</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>+0.4</td>
</tr>
<tr>
<td>Figure</td>
<td>+0.2</td>
</tr>
<tr>
<td>Formation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Prepositions</td>
<td>+0.2</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>+0.2</td>
</tr>
<tr>
<td>Pronouns</td>
<td>+0.6</td>
</tr>
<tr>
<td>Negation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Comprehension</td>
<td>+0.2</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>+3.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>0</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>+0.8</td>
</tr>
<tr>
<td>Figure</td>
<td>+0.2</td>
</tr>
<tr>
<td>Formation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Prepositions</td>
<td>+0.2</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>+0.2</td>
</tr>
<tr>
<td>Pronouns</td>
<td>+0.6</td>
</tr>
<tr>
<td>Negation</td>
<td>+0.2</td>
</tr>
<tr>
<td>Comprehension</td>
<td>+0.2</td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>+3.0</td>
</tr>
</tbody>
</table>
This distinction is of relevance because it is quite possible for a test to be relatively valid for the measurement of one kind of characteristic but entirely invalid for the assessment of other characteristics. Thus it is not meaningful to discuss the validity of the measuring instrument itself, but that of the measuring instrument in relation to the purpose for which it is designed, or for which it is being used.

While the definition of test validity seems simple, there are several types of validity which are of concern to test users. Each of these types of validity takes a somewhat different approach to the assessment of the extent to which a test measures what it is intended to measure. Details concerning the different types of validity and their meaning, usage, and limitations are discussed in text books concerned with educational measurement.

Here the five most basic types of validity will be described. These are predictive validity (frequently referred to as criterion-related validity), concurrent validity, content validity, face validity, and construct validity.

Nunnally (1978) suggests that predictive validity 'is at issue when the purpose is to use an instrument to estimate some important form of behaviour that is external to the measuring instrument itself, the latter being referred to as the criterion' p.87. For example, one validates a university entrance test by showing the accuracy with which it predicts how well a group of students will perform in their academic studies at university.

In operational terms, the degree of criterion-related or predictive validity of the test is estimated by the size of the correlation between test performance and performance on the criterion variable. This correlation coefficient tends to be referred to as a validity coefficient. It is obvious that a test will not be a valid predictor of a criterion unless it correlates substantially with the criterion variable and, similarly, the higher this correlation the more valid a predictor is the test expected to be for the particular criterion.

The term concurrent validity is used when the criterion measure against which test scores are validated is obtained at approximately the same time as the test scores. The American Psychological Association Standards for Educational and Psychological Tests (APA, 1974) differentiates between predictive and concurrent validity only on the basis of the time which elapses between the assessment of the criterion and the obtaining of test results.

Major limitations of criterion validation procedures are that in many situations the relevance and the measurement of the criterion itself may be questionable. In many situations, too, especially when dealing with more abstract developmental concepts, there are no appropriate criteria against which a test can be validly evaluated.

Criterion and predictive validity have their major relevance in relation to selection tests. The tests contained in the ACER Early School Series are designed as diagnostic tests. Their purpose is not to predict future performance, but rather to describe the individuals' present developmental status, and to provide the teacher with a basis for individualized educational planning.

Content validity depends on the extent to which the test reflects a specified domain of content. A thorough examination of the literature provided an
Table 3.1 Multi-skill Multi-sample Matrices for the Auditory Analysis and Conceptual Skills Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Sample 1</th>
<th></th>
<th>Sample 2</th>
<th></th>
<th>Sample 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rec. Con. Audit. Num. 1 Fig. Form.</td>
<td>Rec. Con. Audit. Num. 2 Fig. Form.</td>
<td>Rec. Con. Audit. Num. 3 Fig. Form.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 1</td>
<td>0.49</td>
<td>0.36</td>
<td>0.28</td>
<td></td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>Rec. Con.</td>
<td>0.51</td>
<td>0.15</td>
<td>0.28</td>
<td>0.93</td>
<td>0.64</td>
<td>0.86</td>
</tr>
<tr>
<td>Audit. 1</td>
<td>0.32</td>
<td>0.15</td>
<td>0.28</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Num. 1</td>
<td>0.32</td>
<td>0.15</td>
<td>0.28</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Fig. Form. 1</td>
<td>0.32</td>
<td>0.15</td>
<td>0.28</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Sample 2</td>
<td>0.87</td>
<td>0.77</td>
<td>0.87</td>
<td>0.94</td>
<td>0.56</td>
<td>0.14</td>
</tr>
<tr>
<td>Rec. Con.</td>
<td>0.37</td>
<td></td>
<td>0.62</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Audit. 2</td>
<td>0.56</td>
<td></td>
<td>0.62</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Num. 2</td>
<td>0.47</td>
<td></td>
<td>0.62</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Fig. Form. 2</td>
<td>0.47</td>
<td></td>
<td>0.62</td>
<td>0.93</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Sample 3</td>
<td>0.93</td>
<td>0.64</td>
<td>0.88</td>
<td>0.95</td>
<td>0.55</td>
<td>0.40</td>
</tr>
<tr>
<td>Rec. Con.</td>
<td>0.86</td>
<td>0.64</td>
<td>0.88</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audit. 3</td>
<td>0.86</td>
<td>0.64</td>
<td>0.88</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num. 3</td>
<td>0.86</td>
<td>0.64</td>
<td>0.88</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fig. Form. 3</td>
<td>0.86</td>
<td>0.64</td>
<td>0.88</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
indication of important conceptual pre-reading and language skills. The items contained in the tests of the ACER Early School Series are intended to reflect the content of early skill development. There are no well-defined objective criteria of content validity. ‘Inevitably content validity rests mainly on appeals to reason regarding the adequacy with which important content has been sampled and on the adequacy with which the content has been cast in the form of test items’ (Nunnally, 1978, p.93).

Because the ACER Early School Series is designed to provide a means of identifying, early in their school career, children who for any number of reasons might be expected to experience problems in learning to read or other school related activities, and because the individual tests are intended to provide diagnostic information concerning specific areas and starting points for intervention, content validity is particularly important. Part II of this Handbook contains a detailed account of the rationale and aims for each test, the selection of items, and their scaling. Careful item construction and test development procedures were employed to ensure that a relevant sample of learning-related skills were incorporated in the ACER Early School Series tests.

Face validity is not a validity in the measurement sense. Anastasi (1976) described it as follows:

It refers not to what the test actually measures, but to what it appears superficially to measure. Face validity pertains to whether the test ‘looks valid’ to the examinees who take it, the administrative personnel who decide on its use, and other technical untrained observers. (p.139)

The manner of presentation of tasks, the type of drawings used, and the story type framework in which most of the items are couched should maximize the child’s interest and motivation. The latter factors contribute considerably to the validity of the test scores obtained by testees; that is, they provide face validity.

Construct validity ‘attempts, at the same time, to verify both the existence of some hypothetical construct, and a given test’s ability to measure that construct’ (Popham, 1975, p.153). This differs from content validity, which judges whether a test measures the variables or behaviours which make up the domain under scrutiny. The importance of construct validity in relation to a test depends on how abstract the function or trait is, which the test purports to measure. In the case of most of the tests described in this Handbook, there is an obvious body of content which describes the characteristics of the domain the test purports to measure. The tests of syntactic structures probably represent the best examples of this. It is not difficult to establish content validity for these tests. The constructs measured by the tests of auditory analysis skills and the conceptual skills tests, though of broader content than the language tests, can still be judged in terms of content validity. Some psychological tests measure constructs for which no obvious body of content is available, for example, creativity, intelligence, anxiety, etc. Indices of construct validity become very important for tests of this kind.

Popham’s definition of construct validity, which was quoted above, stresses the dual function of this type of validity. While it may be unnecessary to examine the existence of a hypothetical construct in criterion or domain referenced tests, it is
Table 3.2 Multi-skill Multi-sample Matrices for the Tests of Syntactic Structures

<table>
<thead>
<tr>
<th>Test</th>
<th>Sample A (English)</th>
<th>Sample B (Non-English)</th>
<th>Sample C (Mixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prep. 1 Verb. 1 Pron. 1 Neg. 1 Com. 1</td>
<td>Prep. 2  Verb. 2 Pron. 2 Neg. 2 Com. 2</td>
<td>Prep. 3 Verb. 3 Pron. 3 Neg. 3 Com. 3</td>
</tr>
<tr>
<td>Sample A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rec. Con.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prep. 1</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron. 1</td>
<td>.69 .74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg. 1</td>
<td>.12 .67 .54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Com. 1</td>
<td>.60 .58 .56 .48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rec. Con.</td>
<td>.87 .98</td>
<td>.48 .49 .51</td>
<td></td>
</tr>
<tr>
<td>Prep. 2</td>
<td>.98 .97 .90</td>
<td>.50 .44 .07</td>
<td></td>
</tr>
<tr>
<td>Pron. 2</td>
<td></td>
<td>.23 .26 .08 .30</td>
<td></td>
</tr>
<tr>
<td>Neg. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Com. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample C</td>
<td>.98 .99 .93 .97 .96</td>
<td>.88 .99 .97 .96 .96</td>
<td>.28 .46 .62 .30 .36  .54</td>
</tr>
<tr>
<td>Prep. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pron. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Com. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
certainly necessary to judge construct validity in relation to its second function; that is, to judge how well the test measures the construct.

Before one can investigate the relationship between a particular skill or trait and other skills or behaviours, one must have confidence in one's measurement of that skill or trait.

The primary concern of users and potential users of the ACER Early School Series tests will be with the adequacy of the tests as measures of particular skills, rather than with the adequacy of any traits which might be determined by the confirmation of certain theoretically predicted associations with measures of other traits and theoretical constructs. Because of this it was regarded as unnecessary, and to some degree inappropriate, to establish construct validity, as traditionally done in norm-referenced tests, by means of factor loadings.

The high stability of the Rasch calibrated item difficulties and ability estimates across different, completely independent samples provides a considerable amount of support for the construct validity of the tests.

Further, an attempt was made to examine the consistency with which the tests measure the skills and constructs they were designed to measure by a partial investigation of the convergent and discriminant validity of the tests (Allen and Yen, 1979; Campbell and Fiske, 1959; Sullivan and Feldman, 1979). Campbell and Fiske (1959) first advocated this method of test validation as a logical extension of the concept of construct validity (APA, 1954) and of Cronbach and Meehl's (1955) theoretical considerations. The method is used when two or more traits are measured by two or more methods.

As used by Campbell and Fiske (1959) the term 'convergent validity' refers to the notion that, if a test provides a valid measure of a trait, then, even if one attempts to measure the trait by quite different methods, the methods ought to produce similar results. In other words, different means of measurement would converge in the measurement of the same trait. Discriminant validity refers to their notion that, if the measurement is valid, its result:

should discriminate traits that are distinct. That is, even if we measure different traits by the same method, our results should not correlate too highly because they are different traits, validly measured ... Thus the correlation across different traits should not be too large or we might begin to suspect that our measures are invalid perhaps due to some sort of methodological contamination. (Sullivan and Feldman, 1979, p.18-19)

The method of convergent- and discriminant validation utilizes a matrix of intercorrelations among measurements representing at least two traits, each measured by at least two methods. Measures of the same trait across methods are expected to correlate higher with each other than they do with different traits involving the different methods, and ideally even within the same method.

Campbell and Fiske's (1959) methodology was used partially in the validation of the present series of tests. Specifically for the present purposes the term 'convergent validity' refers to the extent to which items measuring a particular skill yield similar results in different samples of children, making the fundamental assumption that different and independent samples of children in whom the same skill is measured should converge on similar results. Discriminant validity, however, refers to the extent to which the measurements of different skills within
one sample lead to different results. In other words, discriminant validity implies that the skills which differ from one another should produce different results (i.e. lower correlations) even if they were measured in the same person.

In the investigation of convergent and discriminant validity of the tests in the ACER Early School Series, therefore, measures of the same skills across different samples are expected to correlate more highly than they do with measures of different, though at times related, skills within the same sample.

The performances of three samples of children were intercorrelated for the non-language and language tests separately. The results are presented in Tables 3.1 and 3.2 respectively. The three samples whose results are presented in Table 3.1 were independent and unselected. In Table 3.2, Sample A consisted of children who were native speakers of English. All children in Sample B came from families where English was not spoken in the home, and Sample C was an unselected sample containing children from English and non-English language backgrounds.

The correlations presented within the triangles of these tables represent the relationship between the measurements of different skills within each one of the three samples. The correlation coefficients above the empty triangles reflect the similarities of measures of the same skill in different samples, and hence the convergent validity of the skills measured by the tests.

Inspection of Tables 3.1 and 3.2 shows that all values of the correlation coefficients in the diagonals above the empty triangles are considerably higher than those presented within the triangles. The correlations representing same skill/different sample relationships ranged from .64 to .95 for the auditory analysis and conceptual skills tests, and from .87 to .99 in the tests of language skills. These validities are extremely high. The relatively higher correlations obtained in the language tests can be interpreted to show higher validity of these tests than of the non-language tests. Auditory analysis and conceptual skills are from less easily defined and probably broader domains than the knowledge of syntactic structures assessed in the language tests.
PART II
THE TESTS
Area A: Tests of Auditory Analysis Skills

Purpose
An ability to discriminate between speech sounds is a prerequisite not only for the production of speech (Kamil and Rudegeair, 1972) but also for reading (Bateman, 1968; Robinson, 1972; Rosner and Simon, 1971; Zigmond, 1969).

Both the Auditorily Discrimination Test and the Recognition of Initial Consonant Sounds Test are designed for the assessment of skills of auditory analysis. The term ‘auditory analysis’, as used here, refers to the resolution of spoken words into their phonemic elements. The importance of this skill is stressed by Rosner (1973):

Learning to read, regardless of instructional system, requires the transformation of visual symbols into verbal language, be it audible or covert ... the child must learn to associate the basic coding units of reading ... and blended sounds of each word. (p.60)

The aim of the Auditory Analysis Test contained in the ACER Early School Series is to assess the child’s ability to identify acoustic elements in the spoken word. This skill is regarded as an important preliminary step to the process of recognizing the acoustic components of the reading and spelling code as heard in spoken language.
AUDITORY DISCRIMINATION TEST

Rationale

The term 'auditory discrimination' is generally used to describe the ability to differentiate each sound of the language from every other (e.g. Lowry, 1970; Sabatino and Miller, 1979). As a child develops skill in discriminating sounds, he or she must also become increasingly proficient in recognizing them in a variety of phonetic contexts, and eventually gain the same meaning from the spoken and written word. The child who has difficulty hearing differences in speech sounds also has difficulty repeating sequences involving those sounds.

Some children have difficulty in sound recognition, or the selection of relevant from irrelevant auditory stimuli and such children can be expected to experience difficulty in listening in the classroom. Their behaviour is often marked by distractibility, short attention span, and a lack of response. A high correlation has also been found between poor articulation ability and poor discrimination skills in school beginners (Kamil and Rudegeair, 1972; Kamil, 1970; Skeel, Calfee, and Venezky, 1969).

Some authors (e.g. Hillerich, 1974; Kirk, 1966; McLean, Yoder, and Schiefelbusch, 1972) believe that the development of auditory discrimination skill is strongly related to age, and that four- to five-year-old children are able to hear differences of one phoneme in words unless their hearing or other auditory perceptual mechanisms are physically impaired. Wepman Test users would tend to disagree with this view. Wepman (1958) found that five- and six-year-old children with low scores in auditory discrimination skill developed adequate ability in the subsequent two years. They were thus children whose discriminatory abilities developed relatively slowly. However, the same study suggests that older children with similarly low scores in this ability are less likely to show extensive improvement.

In any event, whether the child's disability is caused by delayed development, physical factors, psychological variables, or all of these, it is advisable for the teacher to become aware of the extent of the problem and, where possible to provide special attention. Such individual differences—even where they are developmentally determined and therefore relatively temporary, coinciding with
the beginning of schooling—can lead to severe and lasting disadvantage for some individuals (e.g. Hammill and Larsen, 1974; Skeel, Calfee, and Venezky, 1969).

Difficulty or inability to perform well on auditory discrimination tasks can be the result of differences in customary sound patterns between languages. This may frequently be the case among school beginners in Australia. On the other hand it can be argued that disabilities, whether in auditory discrimination abilities or in other areas, should be identified even where they may be the result of a physical deficit. The teacher is in an excellent position to alert parents and school health personnel who might further investigate the possibility of such a deficit.

A test of auditory discrimination skills was included in the ACER Early School Series because of the consistently demonstrated importance of this skill for the development of speech, reading, and spelling, and because such a high proportion of children in Australian schools come from non-English-speaking homes.

**Description**

The *Auditory Discrimination Test* is a 40-item test. Each item poses an English speech-sound discrimination task in which the child is required to respond to a stimulus word by indicating which of two pictures, representing similar sounding words, represents the stimulus.

The test includes the same types of sound contrasts as the Wepman Test (1958) yet, because of a different procedure of testing, avoids the two major problems which are frequently encountered in the use of the Wepman Test. It has been found that the two major problems with the Wepman Auditory Discrimination Tests are that young children, and children whose native language is not English, do not always understand the concepts ‘same’ and ‘different’ — auditory memory is a major factor being tested in the Wepman Test (Flower, 1968) — and that it is impossible to separate auditory discrimination and auditory memory processes from the child’s performance on the test.

The same/different problem was avoided in the ACER test by requiring the child to mark the appropriate picture. The contamination of the individual’s performance with auditory memory skills was avoided by providing one stimulus word instead of two. Coltheart and Curthoys (1968) and others have commented on the procedure, used in a number of screening tests, of having only two response alternatives. It is accepted that the testee has a 50 per cent chance of producing the correct response by random guessing, although research has shown that the probability of chance success in a two-picture auditory discrimination task with young children is much less than 50 per cent (Fraser, Bellugi, and Brown, 1963).

The provision of multiple alternatives would certainly increase the memory requirement of the task. In the ACER Auditory Discrimination Test, an attempt was made to eliminate the memory variable as far as possible. Other constraints imposed on the design of the test were the need to reduce the time required for testing and the need to limit financial expenditure. The cost of the test would have been considerably higher had three or four response alternatives been provided. Shortening of the test in favour of additional response alternatives was regarded as undesirable because initial pilot data showed varying difficulty levels for different sound combinations.
As testing is likely to be taking place in very small groups, it was assumed that the tester would be able to identify random marking more readily than would be the case if larger groups were tested. Furthermore it is expected that diagnostic interpretations of the test performances will be made only after taking into account many other sources of inconsistency and disturbance which may have influenced the test results. A further discussion of these is presented at the end of this chapter.

Development

Initial item trials were conducted with three sets of 30 items in a number of kindergartens and inner-city schools. Trials were conducted with short tests of varying item number with two, three, and four response alternatives. The increased memory requirement for a four-choice test caused obvious interference in the performance of some kindergarten and migrant children. A tendency to inspect only the first two or middle two pictures was noted among these children. In the three-choice case, also, the strategy of comparing only two pictures was observed but less frequently.

Different testing procedures were tried out in one-to-one testing situations. For the majority of children, no difference was found between a two stimulus words same/different procedure and the procedure finally adopted in the present test. This procedure was preferred because the same/different procedure complicates response requirements somewhat when used as a paper and pencil test, and because of possible difficulties with the same/different concept (Flower, 1968) especially for very young and migrant children. Other reasons for the choice of the mode of administration adopted in the test were presented in the previous section.

Administration and Scoring

The time required for administration of all 40 items depends on the maturity and possible handicaps of the testee, and the size of the group of individuals to whom the test is being administered. During trial testing, the initial instructions and the administrations of the three practice items took two to five minutes. Individual items tended to require less than 30 seconds each.

The tests included in the ACER Early School Series are designed to be power tests, not speed tests. On the basis of the experience gained from trial testing, it is suggested that the performance of very slow responders is unlikely to be improved if more than 30 seconds are allowed for each item.

The Auditory Discrimination Test can be administered in parts over a number of sessions. In this case, the three practice items are administered at the beginning of each testing session, and they are followed by the item subsequent to the last item administered during the previous session. This procedure is recommended particularly for very young and handicapped children. Materials required consist of the Directions for Administration which contain general and detailed instructions for the administration of the Auditory Discrimination Test, a test booklet, cardboard marker, crayon or pencil for each child, and a spare test booklet which the tester uses for demonstration purposes. The instructions are in simple English and are printed out in full for each item.
### Table 4.1 Auditory Discrimination Test: Performance of 194 School Beginners

<table>
<thead>
<tr>
<th></th>
<th>Home language</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total sample</td>
<td>Non-English</td>
<td>English</td>
</tr>
<tr>
<td>Mode</td>
<td>40 (54%)</td>
<td>40 (39%)</td>
<td>40 (57%)</td>
</tr>
<tr>
<td>Median</td>
<td>40</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>Mean</td>
<td>38.6</td>
<td>37.0</td>
<td>39.0</td>
</tr>
<tr>
<td>SD</td>
<td>3.7</td>
<td>4.0</td>
<td>3.5</td>
</tr>
<tr>
<td>SE</td>
<td>0.3</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Average item facility</td>
<td>96.3%</td>
<td>92.5%</td>
<td>97.4%</td>
</tr>
<tr>
<td>SD</td>
<td>3.0%</td>
<td>8.0%</td>
<td>2.2%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.92</td>
<td>0.87</td>
<td>0.94</td>
</tr>
<tr>
<td>N</td>
<td>194</td>
<td>38</td>
<td>156</td>
</tr>
</tbody>
</table>

### Scoring

A score key is provided. One point is given for each correct item. Possible points for each item are zero (0) or one (1). The raw score for the test is provided by the sum of correctly completed items. Possible total scores for the test range from zero (0) to forty (40). Tables 4.4 and A.2 allow the conversion of raw scores to Rasch scaled developmental ability scores.

### Research Results

This section contains descriptions of the performances of different samples of children on the Auditory Discrimination Test. These research results are presented here to assist users of the test in their interpretation of test results by providing them with comparative data.

Table 4.1 presents a summary of the performances of 194 children from the representative sample of 197 school beginners, described in Chapter 3, who were tested approximately six weeks after the beginning of the school year.

As shown in Table 4.1, among the total sample, 54 per cent of children completed all items correctly, i.e., obtained a total score of 40 on the test. This performance level is lower than might have been expected in a perceptual ability which is generally expected to have developed adequately at the stage when the child begins formal schooling. Only 39 per cent of the children from non-English language backgrounds completed all items correctly. This information is presented here, not to underline differences between groups of children, but rather to focus on the fact that a large proportion of children have not reached a level of auditory discrimination ability, which many teachers and parents tend to take for granted. Criterion cut-off scores are naturally inappropriate in a diagnostic test of this type.

To validate further the findings obtained from the representative sample, the results of another 599 school beginners were provided by 13 teachers and school counsellors who tested unselected samples from city and rural schools of varying sizes. For analysis, these data were divided into two sets according to the expected proportion of non-English language background children in the sample schools.
As expected, the Auditory Discrimination Test results of the two samples differed to the extent that in the sample including fewer children from non-English language backgrounds, which consisted of 249 school beginners, 57 per cent of children obtained full score on the test, while in the other sample, which consisted of 359 school beginners, only 27 per cent completed the test without errors. Testing of both samples took place in the same week, eight weeks after the beginning of the school year.

Table 4.2 shows the difficulty found by children in the different samples for each item of the Auditory Discrimination Test. The facility of each item is expressed in terms of the percentage of children in the sample who completed the item correctly. The ordering of items in the final version of the test was determined by the performance of the total representative sample. The item difficulties for the validity samples are presented in the last two columns of Table 4.2. The headings A and B refer to the samples with a high and low proportion of non-English language background children respectively.

Table 4.2 shows that for most items only a small percentage of children experienced difficulty. This percentage was considerably higher among children from non-English backgrounds than among the native speakers of English. An appreciable difference in the within-group variability of difficulties is reflected in the observable difference in standard deviations. The native speakers of English tended to do better on the test, and the variation between individuals within this group was considerably smaller than among non-English language background children. It is interesting to note that, although the latter tended to perform less well on the test as a whole than their English language background peers, the former performed better on some individual items. This emphasizes the importance of a focus on individual differences and the diagnosis of strengths and weaknesses of individuals, rather than a differentiation of one group of children from another or a categorization and labelling of children.

Interpretation of Scores
Table 4.3 shows the estimates of Rasch item difficulties (Di) and standard errors (SEDi) for the difficulty estimate for each item in the representative and the validation samples. No significant difference was found in the Rasch item difficulties between the three samples (F = 0.01, df = 2, 117, p > 0.05).

Table 4.4 shows Rasch-scaled ability estimates (Ai) and standard errors (SEAi) for each estimate for the three samples. Inspection of Table 4.4 indicates that the Rasch ability estimates obtained from raw scores in the three separate and independent samples show very close agreement. This shows the stability of the Auditory Discrimination Test as a measurement tool. The Rasch-scaled ability estimates are quite independent of the sample of children from which they were obtained. In other words the Rasch ability obtained by a child will be the same no matter in which group of children the child is tested.

The correlations between the item difficulties obtained for the three completely independent groups were positive and reasonably high. This suggests that, while the ease with which different groups solved the tasks varied because of developmental or other differences between these groups, the performances of the
### Table 4.2 Auditory Discrimination Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Non-English</td>
<td>A</td>
</tr>
<tr>
<td>1 bee</td>
<td>99</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>2 soap</td>
<td>99</td>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>3 bed</td>
<td>99</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>4 bat</td>
<td>99</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>5 sauce</td>
<td>99</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>6 puzzle</td>
<td>99</td>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>7 tree</td>
<td>99</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>8 ship</td>
<td>99</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>9 feather</td>
<td>99</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>10 see saw</td>
<td>99</td>
<td>100</td>
<td>95</td>
</tr>
<tr>
<td>11 fish</td>
<td>98</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>12 bell</td>
<td>98</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>13 clock</td>
<td>98</td>
<td>100</td>
<td>93</td>
</tr>
<tr>
<td>14 money</td>
<td>98</td>
<td>97</td>
<td>98</td>
</tr>
<tr>
<td>15 rock</td>
<td>98</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>16 plane</td>
<td>98</td>
<td>95</td>
<td>98</td>
</tr>
<tr>
<td>17 grass</td>
<td>98</td>
<td>97</td>
<td>98</td>
</tr>
<tr>
<td>18 cotton</td>
<td>98</td>
<td>97</td>
<td>94</td>
</tr>
<tr>
<td>19 tooth</td>
<td>98</td>
<td>92</td>
<td>93</td>
</tr>
<tr>
<td>20 play</td>
<td>98</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>21 pig</td>
<td>97</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>22 sweep</td>
<td>97</td>
<td>92</td>
<td>96</td>
</tr>
<tr>
<td>23 sheep</td>
<td>97</td>
<td>97</td>
<td>95</td>
</tr>
<tr>
<td>24 toe</td>
<td>97</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>25 pen</td>
<td>96</td>
<td>92</td>
<td>81</td>
</tr>
<tr>
<td>26 wash</td>
<td>96</td>
<td>89</td>
<td>95</td>
</tr>
<tr>
<td>27 roof</td>
<td>96</td>
<td>89</td>
<td>97</td>
</tr>
<tr>
<td>28 pet</td>
<td>95</td>
<td>84</td>
<td>86</td>
</tr>
<tr>
<td>29 fruit</td>
<td>95</td>
<td>92</td>
<td>94</td>
</tr>
<tr>
<td>30 stool</td>
<td>95</td>
<td>84</td>
<td>93</td>
</tr>
<tr>
<td>31 pin</td>
<td>95</td>
<td>89</td>
<td>94</td>
</tr>
<tr>
<td>32 path</td>
<td>94</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td>33 town</td>
<td>94</td>
<td>87</td>
<td>87</td>
</tr>
<tr>
<td>34 pearl</td>
<td>93</td>
<td>87</td>
<td>90</td>
</tr>
<tr>
<td>35 mouth</td>
<td>93</td>
<td>76</td>
<td>90</td>
</tr>
<tr>
<td>36 peel</td>
<td>93</td>
<td>92</td>
<td>88</td>
</tr>
<tr>
<td>37 knot</td>
<td>92</td>
<td>82</td>
<td>82</td>
</tr>
<tr>
<td>38 board</td>
<td>90</td>
<td>74</td>
<td>88</td>
</tr>
<tr>
<td>39 cap</td>
<td>89</td>
<td>79</td>
<td>66</td>
</tr>
<tr>
<td>40 fountain</td>
<td>87</td>
<td>71</td>
<td>86</td>
</tr>
</tbody>
</table>

| X | 96.3 | 92.5 | 92.8 | 95.8 |  |
| SD | 3.0 | 8.0 | 6.4 | 3.2 |  |
| N | 194 | 38 | 350 | 249 |  |
Table 4.3 Auditory Discrimination Test: Rasch-scaled Items, Difficulty Estimates (Di); and Standard Errors (SE(Di)) for Three Independent Samples

<table>
<thead>
<tr>
<th>Item</th>
<th>Di</th>
<th>SE(Di)</th>
<th>Di</th>
<th>SE(Di)</th>
<th>Di</th>
<th>SE(Di)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61.6</td>
<td>1.3</td>
<td>56.7</td>
<td>0.9</td>
<td>56.4</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>60.7</td>
<td>1.4</td>
<td>64.3</td>
<td>0.7</td>
<td>60.2</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>60.0</td>
<td>1.4</td>
<td>55.6</td>
<td>1.0</td>
<td>54.8</td>
<td>1.6</td>
</tr>
<tr>
<td>4</td>
<td>57.9</td>
<td>1.6</td>
<td>58.6</td>
<td>0.9</td>
<td>58.1</td>
<td>1.3</td>
</tr>
<tr>
<td>5</td>
<td>57.0</td>
<td>1.7</td>
<td>54.1</td>
<td>1.1</td>
<td>55.7</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>57.0</td>
<td>1.7</td>
<td>54.1</td>
<td>1.1</td>
<td>53.8</td>
<td>1.7</td>
</tr>
<tr>
<td>7</td>
<td>57.0</td>
<td>1.7</td>
<td>55.6</td>
<td>1.0</td>
<td>55.7</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>56.4</td>
<td>1.7</td>
<td>57.9</td>
<td>0.9</td>
<td>59.1</td>
<td>1.3</td>
</tr>
<tr>
<td>9</td>
<td>55.2</td>
<td>1.9</td>
<td>55.9</td>
<td>1.0</td>
<td>56.1</td>
<td>1.5</td>
</tr>
<tr>
<td>10</td>
<td>55.2</td>
<td>1.9</td>
<td>56.9</td>
<td>0.9</td>
<td>55.7</td>
<td>1.5</td>
</tr>
<tr>
<td>11</td>
<td>55.2</td>
<td>1.9</td>
<td>50.2</td>
<td>1.4</td>
<td>55.7</td>
<td>1.5</td>
</tr>
<tr>
<td>12</td>
<td>55.2</td>
<td>1.9</td>
<td>51.5</td>
<td>1.3</td>
<td>54.8</td>
<td>1.6</td>
</tr>
<tr>
<td>13</td>
<td>52.8</td>
<td>2.3</td>
<td>50.3</td>
<td>1.4</td>
<td>48.4</td>
<td>2.6</td>
</tr>
<tr>
<td>14</td>
<td>52.8</td>
<td>2.3</td>
<td>59.1</td>
<td>0.8</td>
<td>59.6</td>
<td>1.3</td>
</tr>
<tr>
<td>15</td>
<td>52.8</td>
<td>2.3</td>
<td>49.2</td>
<td>1.5</td>
<td>55.2</td>
<td>1.6</td>
</tr>
<tr>
<td>16</td>
<td>51.7</td>
<td>2.5</td>
<td>46.8</td>
<td>1.8</td>
<td>50.4</td>
<td>2.2</td>
</tr>
<tr>
<td>17</td>
<td>51.7</td>
<td>2.5</td>
<td>46.2</td>
<td>1.9</td>
<td>52.9</td>
<td>2.9</td>
</tr>
<tr>
<td>18</td>
<td>50.5</td>
<td>2.8</td>
<td>43.1</td>
<td>2.4</td>
<td>43.0</td>
<td>4.1</td>
</tr>
<tr>
<td>19</td>
<td>50.4</td>
<td>2.8</td>
<td>47.3</td>
<td>1.8</td>
<td>50.4</td>
<td>2.2</td>
</tr>
<tr>
<td>20</td>
<td>50.5</td>
<td>2.8</td>
<td>49.9</td>
<td>1.5</td>
<td>52.0</td>
<td>2.0</td>
</tr>
<tr>
<td>21</td>
<td>48.9</td>
<td>3.2</td>
<td>51.5</td>
<td>1.3</td>
<td>51.2</td>
<td>2.1</td>
</tr>
<tr>
<td>22</td>
<td>48.9</td>
<td>3.2</td>
<td>47.8</td>
<td>1.7</td>
<td>52.0</td>
<td>2.0</td>
</tr>
<tr>
<td>23</td>
<td>46.6</td>
<td>3.9</td>
<td>47.8</td>
<td>1.7</td>
<td>39.0</td>
<td>5.7</td>
</tr>
<tr>
<td>24</td>
<td>46.6</td>
<td>3.9</td>
<td>44.0</td>
<td>2.3</td>
<td>39.0</td>
<td>5.7</td>
</tr>
<tr>
<td>25</td>
<td>46.6</td>
<td>3.9</td>
<td>51.5</td>
<td>1.3</td>
<td>50.4</td>
<td>2.2</td>
</tr>
<tr>
<td>26</td>
<td>46.6</td>
<td>3.9</td>
<td>43.1</td>
<td>2.5</td>
<td>45.4</td>
<td>3.3</td>
</tr>
<tr>
<td>27</td>
<td>46.6</td>
<td>3.9</td>
<td>45.6</td>
<td>2.0</td>
<td>43.0</td>
<td>4.1</td>
</tr>
<tr>
<td>28</td>
<td>46.6</td>
<td>3.9</td>
<td>44.1</td>
<td>2.3</td>
<td>39.0</td>
<td>5.7</td>
</tr>
<tr>
<td>29</td>
<td>46.6</td>
<td>3.9</td>
<td>44.8</td>
<td>2.1</td>
<td>48.4</td>
<td>2.6</td>
</tr>
<tr>
<td>30</td>
<td>46.6</td>
<td>3.9</td>
<td>50.6</td>
<td>1.4</td>
<td>54.3</td>
<td>1.7</td>
</tr>
<tr>
<td>31</td>
<td>42.8</td>
<td>5.4</td>
<td>46.2</td>
<td>1.9</td>
<td>49.5</td>
<td>2.4</td>
</tr>
<tr>
<td>32</td>
<td>42.8</td>
<td>5.4</td>
<td>47.3</td>
<td>1.8</td>
<td>45.4</td>
<td>3.3</td>
</tr>
<tr>
<td>33</td>
<td>42.8</td>
<td>5.4</td>
<td>48.3</td>
<td>1.6</td>
<td>52.6</td>
<td>1.9</td>
</tr>
<tr>
<td>34</td>
<td>42.8</td>
<td>5.4</td>
<td>38.9</td>
<td>3.5</td>
<td>43.0</td>
<td>4.1</td>
</tr>
<tr>
<td>35</td>
<td>42.8</td>
<td>5.4</td>
<td>45.6</td>
<td>2.0</td>
<td>43.0</td>
<td>4.1</td>
</tr>
<tr>
<td>36</td>
<td>42.8</td>
<td>5.4</td>
<td>51.5</td>
<td>1.3</td>
<td>48.4</td>
<td>2.6</td>
</tr>
<tr>
<td>37</td>
<td>42.8</td>
<td>5.4</td>
<td>49.2</td>
<td>1.5</td>
<td>47.1</td>
<td>2.9</td>
</tr>
<tr>
<td>38</td>
<td>42.8</td>
<td>5.4</td>
<td>46.2</td>
<td>1.9</td>
<td>45.4</td>
<td>3.3</td>
</tr>
<tr>
<td>39</td>
<td>42.8</td>
<td>5.4</td>
<td>48.7</td>
<td>1.6</td>
<td>43.0</td>
<td>4.1</td>
</tr>
<tr>
<td>40</td>
<td>42.8</td>
<td>5.4</td>
<td>44.0</td>
<td>2.3</td>
<td>39.0</td>
<td>5.7</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td></td>
<td>50.0</td>
<td></td>
<td>50.1</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>5.8</td>
<td></td>
<td>5.3</td>
<td></td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>89</td>
<td></td>
<td>254</td>
<td></td>
<td>106</td>
<td></td>
</tr>
</tbody>
</table>

a The differences in N between Tables 4.1 and 4.2, and Tables 4.3 and 4.4 are due to the fact that Rasch analyses exclude individuals who have obtained full scores or zero scores on the test.
### THE TESTS

#### Table 4.4: Auditory Discrimination Test: Rasch Ability Estimates (\(A_i\)) and Standard Errors (\(SE_{(A_i)}\)) for Three Independent Samples

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Representative sample</th>
<th>Validation samples</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A_i)</td>
<td>(SE_{(A_i)})</td>
<td>(A_i)</td>
<td>(SE_{(A_i)})</td>
<td>(A_i)</td>
</tr>
<tr>
<td>1</td>
<td>29.1</td>
<td>5.8</td>
<td>29.7</td>
<td>5.6</td>
<td>28.7</td>
</tr>
<tr>
<td>2</td>
<td>33.2</td>
<td>4.1</td>
<td>33.7</td>
<td>4.0</td>
<td>32.9</td>
</tr>
<tr>
<td>3</td>
<td>35.6</td>
<td>3.4</td>
<td>36.1</td>
<td>3.3</td>
<td>35.4</td>
</tr>
<tr>
<td>4</td>
<td>37.4</td>
<td>3.0</td>
<td>37.8</td>
<td>2.9</td>
<td>37.2</td>
</tr>
<tr>
<td>5</td>
<td>38.9</td>
<td>2.7</td>
<td>39.2</td>
<td>2.7</td>
<td>38.7</td>
</tr>
<tr>
<td>6</td>
<td>40.6</td>
<td>2.5</td>
<td>40.4</td>
<td>2.5</td>
<td>39.9</td>
</tr>
<tr>
<td>7</td>
<td>41.1</td>
<td>2.4</td>
<td>41.4</td>
<td>2.3</td>
<td>41.0</td>
</tr>
<tr>
<td>8</td>
<td>42.1</td>
<td>2.3</td>
<td>42.3</td>
<td>2.2</td>
<td>41.9</td>
</tr>
<tr>
<td>9</td>
<td>42.9</td>
<td>2.2</td>
<td>43.1</td>
<td>2.1</td>
<td>42.8</td>
</tr>
<tr>
<td>10</td>
<td>43.7</td>
<td>2.1</td>
<td>43.9</td>
<td>2.0</td>
<td>43.6</td>
</tr>
<tr>
<td>11</td>
<td>44.5</td>
<td>2.0</td>
<td>44.6</td>
<td>2.0</td>
<td>44.4</td>
</tr>
<tr>
<td>12</td>
<td>45.2</td>
<td>2.0</td>
<td>45.3</td>
<td>1.9</td>
<td>45.1</td>
</tr>
<tr>
<td>13</td>
<td>45.8</td>
<td>1.9</td>
<td>45.9</td>
<td>1.9</td>
<td>54.3</td>
</tr>
<tr>
<td>14</td>
<td>46.5</td>
<td>1.9</td>
<td>46.6</td>
<td>1.8</td>
<td>46.4</td>
</tr>
<tr>
<td>15</td>
<td>47.1</td>
<td>1.9</td>
<td>47.2</td>
<td>1.8</td>
<td>47.0</td>
</tr>
<tr>
<td>16</td>
<td>47.7</td>
<td>1.8</td>
<td>47.8</td>
<td>1.8</td>
<td>47.6</td>
</tr>
<tr>
<td>17</td>
<td>48.3</td>
<td>1.8</td>
<td>48.2</td>
<td>1.8</td>
<td>48.2</td>
</tr>
<tr>
<td>18</td>
<td>48.9</td>
<td>1.8</td>
<td>48.9</td>
<td>1.8</td>
<td>48.8</td>
</tr>
<tr>
<td>19</td>
<td>49.4</td>
<td>1.8</td>
<td>49.4</td>
<td>1.8</td>
<td>49.4</td>
</tr>
<tr>
<td>20</td>
<td>50.0</td>
<td>1.8</td>
<td>50.0</td>
<td>1.8</td>
<td>50.0</td>
</tr>
<tr>
<td>21</td>
<td>50.6</td>
<td>1.8</td>
<td>50.6</td>
<td>1.8</td>
<td>50.6</td>
</tr>
<tr>
<td>22</td>
<td>51.1</td>
<td>1.8</td>
<td>51.1</td>
<td>1.8</td>
<td>51.2</td>
</tr>
<tr>
<td>23</td>
<td>51.7</td>
<td>1.8</td>
<td>51.7</td>
<td>1.8</td>
<td>51.8</td>
</tr>
<tr>
<td>24</td>
<td>52.3</td>
<td>1.8</td>
<td>52.2</td>
<td>1.8</td>
<td>52.4</td>
</tr>
<tr>
<td>25</td>
<td>52.9</td>
<td>1.9</td>
<td>52.8</td>
<td>1.8</td>
<td>53.0</td>
</tr>
<tr>
<td>26</td>
<td>53.5</td>
<td>1.9</td>
<td>53.4</td>
<td>1.8</td>
<td>53.6</td>
</tr>
<tr>
<td>27</td>
<td>54.2</td>
<td>1.9</td>
<td>54.1</td>
<td>1.9</td>
<td>54.3</td>
</tr>
<tr>
<td>28</td>
<td>54.8</td>
<td>2.0</td>
<td>54.7</td>
<td>1.9</td>
<td>54.9</td>
</tr>
<tr>
<td>29</td>
<td>55.5</td>
<td>2.0</td>
<td>55.4</td>
<td>2.0</td>
<td>55.7</td>
</tr>
<tr>
<td>30</td>
<td>56.3</td>
<td>2.1</td>
<td>56.1</td>
<td>2.0</td>
<td>56.4</td>
</tr>
<tr>
<td>31</td>
<td>57.1</td>
<td>2.2</td>
<td>56.9</td>
<td>2.1</td>
<td>57.2</td>
</tr>
<tr>
<td>32</td>
<td>57.9</td>
<td>2.3</td>
<td>57.7</td>
<td>2.2</td>
<td>58.1</td>
</tr>
<tr>
<td>33</td>
<td>58.9</td>
<td>2.4</td>
<td>58.6</td>
<td>2.3</td>
<td>59.0</td>
</tr>
<tr>
<td>34</td>
<td>59.9</td>
<td>2.5</td>
<td>59.6</td>
<td>2.5</td>
<td>60.1</td>
</tr>
<tr>
<td>35</td>
<td>61.1</td>
<td>2.7</td>
<td>60.8</td>
<td>2.7</td>
<td>61.3</td>
</tr>
<tr>
<td>36</td>
<td>62.6</td>
<td>3.0</td>
<td>62.2</td>
<td>2.9</td>
<td>62.8</td>
</tr>
<tr>
<td>37</td>
<td>64.4</td>
<td>3.4</td>
<td>63.9</td>
<td>3.3</td>
<td>64.6</td>
</tr>
<tr>
<td>38</td>
<td>66.8</td>
<td>4.1</td>
<td>66.3</td>
<td>4.0</td>
<td>67.1</td>
</tr>
<tr>
<td>39</td>
<td>70.9</td>
<td>5.8</td>
<td>70.3</td>
<td>5.6</td>
<td>71.3</td>
</tr>
</tbody>
</table>

\(N = 89\)
representative sample and validation sample A showed reasonable similarity in the rank order of item difficulties. Individuals in validation sample B appeared to differ more from the other two groups.

The ordering of the ability estimates obtained in the three samples was very similar and shows the stability of the measurement as it relates to the locations of particular individuals on the ability assessment provided by the test. The ability estimate provided by the Auditory Discrimination Test is thus shown to be unaffected by the performance of the group within which the individual is tested.

Table 4.5 shows the correlations between Rasch item difficulties on the left side, and the correlations between the ability estimates for the groups on the right side.

Implications
Although the ability to discriminate between sounds, like any other perceptual ability, is considered to be developmental in nature and, although some younger or more slowly developing children have been observed to make rapid improvement in these skills when they grew older, it cannot be denied that the very fact that some children begin school at a stage when their ability to discriminate sounds is below the level of that of their peers and/or below a level of adequacy indicates a need for intervention and remediation.

Compensatory training is particularly important in the case of handicapped children. For individuals who, in a number of retest situations, show little or no improvement in auditory discrimination ability, full consideration needs to be given to this problem in the design of language and reading programs. For example, Johnson and Myklebust (1967) showed that many aphasic children have severe auditory discrimination difficulties along with delayed language development. For these children and children with physiologically based hearing deficits, compensatory teaching strategies in language and initial reading instruction should certainly make greater use of visual and kinesthetic cues. The literature, particularly in the ERIC system, contains many references which will be found useful by teachers both in regular and special classrooms, and by parents and others who are engaged in helping children overcome delayed development or to compensate for deficits in this area.

Suggestions of specific activities, which may be found useful in the training of auditory analysis skills more generally, are provided at the end of Chapter 5, as they cover the skills assessed by both the Auditory Discrimination Test and the Recognition of Initial Consonant Sounds Test.

It is important to consider certain limitations of the test. Tests of auditory discrimination provide a useful procedure to determine whether an individual can distinguish sound similarities between syllables or words, but they do not identify individuals whose primary disability relates to sound perception within words. Since many different kinds of auditory interferences and disturbances (e.g. lack of vocabulary, phonetic incompetence, foreign accents, inconsistencies in the tester's presentation of stimulus words, environmental noise, etc.) may be influencing the child's test performance, it is necessary to explore the individual testee's auditory
Table 4.5 Auditory Discrimination Test: Correlations between Item Difficulties and between Ability Estimates for Three Independent Samples of School Beginners

<table>
<thead>
<tr>
<th>Sample</th>
<th>Item difficulty (Di)</th>
<th>Ability estimate (Ai)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Validity A</td>
<td>Validity B</td>
</tr>
<tr>
<td>Representative</td>
<td>.71</td>
<td>.41</td>
</tr>
<tr>
<td>Validity A</td>
<td>.51</td>
<td></td>
</tr>
</tbody>
</table>

ability in many bordering areas. These areas include:

1. hearing similarities and differences in words
2. distinguishing similar parts of words
3. blending sounds into words
4. following auditory sequences
5. analysis and synthesis of syllables and individual sound sequences.
RECOGNITION OF INITIAL CONSONANT SOUNDS TEST

Rationale
The general points made in the purpose statement introducing the tests of auditory analysis skills section of the Handbook and in the rationale for the Auditory Discrimination Test cover the Recognition of Initial Consonant Sounds Test. These points will not be repeated here. The reader intending to use the Recognition of Initial Consonant Sounds Test is advised to consult the relevant sections earlier in this Handbook.

The content of the following paragraphs will be concerned with those aspects of the assessment of auditory analysis skills which are specific to the ability to recognize initial consonants.

Letters in the initial and final positions of words have a special role in reading. They begin or end not only their word but also a phoneme, a syllable, or morpheme, and usually a spelling pattern. (Ruben, 1978, p.171)

Research has shown that children just learning to read use initial letters as cues in word recognition more often than they use either medial letters or word shape (Marchbanks and Levin, 1965; Rubin, 1978; Williams, Blumberg, and Williams, 1970).

Although adults utilize initial letters less often than children (Williams, Blumberg, and Williams, 1970) it has been shown that they are more likely to perceive initial letters more accurately than letters in any other position in tachistoscopic presentations (Bruner and O'Dowd, 1958; Haslerud and Clark, 1957). Initial letters play an important role in facilitating the recall of words in the absence of visual presentation: On occasions when a person finds himself in the state of having a word ‘on the tip of the tongue’, he or she is most likely to remember initial sounds or letters (Brown and McNeill, 1966; Rubin 1975, 1978). Vowels have been found to be used more frequently in word recognition than consonants (Groff, 1978) and are generally regarded as learnt more easily. For this reason the present test is based on the recognition of initial consonants and does not include words starting with vowel sounds.

The Auditory Discrimination Test and the Word Knowledge Test focus on whole words. The ability to segment words is increasingly being regarded as an important

The processes underlying the recognition of initial consonants include the analysis and synthesis of words. Practice in this type of task can be expected to increase the child's sensitivity to the structure of words (Goldstein, 1976; Otto and Chester, 1976). Goldstein's proposition was that a phonemic sensitivity might operate both as a cause and a consequence of linguistic capability and learning to read; that is, that the possession of these skills helps in the learning of reading but that learning to read also improves these skills.

Goldstein's (1966) findings and those of Fox and Routh (1976) and Liberman and Shankweiler (1977) suggest that increased sensitivity to word and letter sounds may well be operating as a facilitator in the acquisition of language and reading skills.

The skills of phonemic analysis and synthesis taught and assessed by the Recognition of Initial Consonant Sounds Test were found to correlate with learning to read (Calfee, Lindamood, and Lindamood, 1973; Chall, Roswell, and Blumenthal, 1963; Rosner, 1973): As was suggested previously, the auditory analysis skills assessed in both the Auditory Discrimination Test and in the Recognition of Initial Consonant Sounds Test have been regarded as a prerequisite for and a component of the processes of reading and spelling. Both these activities demand that the individual is able to recognize and to relate the phonemic elements of the spoken word to the printed or written code.

Focusing on initial consonants is expected to acquaint the child with a further basic concept, namely that the phonemic elements in a word have a specific temporal organization. In other words, the individual learns that the blended sounds of a spoken word occur in a precise sequence.

In this connection, it is important to recognize that an individual does not need to know the names of letters in order to learn to read. Rather he or she should have established certain frequently occurring sound associations for letters. The latter skill and the ability to name letters are quite different. (An extensive discussion of evidence on this point is provided by Hillerich, 1966.) Tests of knowledge of letter names are a type of test which correlates well with reading success, but this type of test has no diagnostic value.

A further purpose of the Recognition of Initial Consonant Sounds Test materializes from the methodology of test administration. Each consonant sound is taught as part of the item by which it is tested. This procedure was designed not only to utilize testing time for instructional purposes but, more importantly, to provide the teacher with an indication of the child's ability to learn.

Description
The Recognition of Initial Consonant Sounds Test is a 17-item pictorial test. Each item consists of two parts. During the initial part of the item, the tester teaches the relevant initial consonant sound utilizing four pictorial representations of words beginning with that sound or sound combination. The second part of each item
requires the testee to respond to four stimulus words of which two begin with the previously taught consonant sound. The testee's response consists of placing a cross (X) on the correct pictures.

The pictures consist of simple, easily recognizable, single object drawings. Each page of the disposable test booklet contains three items.

**Development**

The initial item pool of 20 items was tried out first in one-to-one testing situations, and later with groups of between three and eight children in two kindergartens and in preparatory grades in three schools. A number of teachers assisted in these trials. Three out of the 20 items were consistently found to be much more difficult than the other items in the test. They were items relating to consonant sound combinations which were, as a result of advice from teachers and panel members, subsequently deleted from the test. This left the Recognition of Initial Consonant Sounds Test as a 17-item test.

Research data on the test were gathered from a total of 797 school beginners. The representative sample, described in Chapter 3, consisted of 197 children. Performance data from four additional, unselected samples were provided by teachers and school counsellors. Two of these samples consisted of 150 children each. The other two groups consisted of 200 and 100 children each. The areas in which the schools constituting these samples were situated differed in socio-economic status and in the proportion of pupils from a non-English language background. Testing of these samples did not necessarily coincide.

**Administration and Scoring**

As noted in the description of this test, the administration of each item is preceded by a preliminary exercise which is intended to teach the consonant sound which is to be tested.

The time required for each item depends on the maturity and possible handicaps of the testee. The amount of school experience and the size of the group of individuals to whom the test is being administered at one time will influence the time required for this test more than any of the other tests contained in the ACER Early School Series.

Unless used as a re-test, for example at the end of the preparatory school year or at the beginning of Year 1, it is suggested that this test be administered in several parts over a number of sessions. When this procedure is followed, the preliminary exercise is administered at the beginning of each testing session. This is then followed by the item subsequent to the last item completed during the previous session.

During trial testing the average time required for the first administration of four items including the practice item to groups of three to five children, early in the preparatory year, was approximately 15 minutes. If the test is administered in the second-half of the school year or on subsequent administrations, this time can be expected to be reduced to approximately 10 minutes. When testing in one-to-one situations, even the first administration of the test should take less than 10 minutes for four items.
Table 5.1 Recognition of Initial Consonant Sounds Test: Performance of 197 School Beginners

<table>
<thead>
<tr>
<th></th>
<th>Obtained scores</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total sample</td>
<td>Home language</td>
</tr>
<tr>
<td></td>
<td>Non-English</td>
<td>English</td>
</tr>
<tr>
<td>Mode</td>
<td>17 (28%)</td>
<td>15 (13%)</td>
</tr>
<tr>
<td>Median</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>12.6</td>
<td>9.3</td>
</tr>
<tr>
<td>SD</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>SE</td>
<td>0.4</td>
<td>-0.9</td>
</tr>
<tr>
<td>Average item facility</td>
<td>74.3%</td>
<td>54.5%</td>
</tr>
<tr>
<td>SD</td>
<td>7.1%</td>
<td>11.5%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.94</td>
<td>0.91</td>
</tr>
<tr>
<td>N</td>
<td>197</td>
<td>37</td>
</tr>
</tbody>
</table>

The tests contained in the ACER Early School Series are power tests, not speed tests. Because of the instructional content of the Recognition of Initial Consonant Sounds Test, speed should, in fact, be largely disregarded.

Materials required consist of the Directions for Administration, a test booklet, cardboard marker, crayon or pencil for each child, and a spare test booklet which the tester can use for demonstration purposes. The Directions for Administration contain full instructions for each item.

Scoring
A score key is provided. A score of one point is given for each correctly completed item. Possible points for each item are zero (0) or one (1). This means that for an item to be completed correctly, both response pictures pertaining to the consonant tested must be masked correctly. The procedure of requiring two responses for each item is designed to increase the reliability of the test. No partial scores are awarded, but the teacher will find the observation of partially correct items of some diagnostic value.

The raw score for the total test is provided by the sum of correctly completed items. Possible total scores for the test range from zero (0) to seventeen (17). Table 5.4 allows for the conversion of raw scores to Rasch-scaled developmental ability scores.

Research Results
This section provides a summary of the results obtained on the Recognition of Initial Consonant Sounds Test by the representative sample described in Chapter 3, and by four validation groups, described in the section on the development of the test.

Table 5.1 describes the performance of the representative sample of 197 school beginners. Inspection of Table 5.1 shows that one third of the children whose home language was English completed all items of the test correctly at its first
administration. Only 13 per cent of the children from non-English language backgrounds were able to complete all items correctly at the initial administration of the test. Comparison of the average performance scores (mode, median, and mean) of the two groups shows, not unexpectedly, an obvious disadvantage in this skill for the children whose home language is not English. This finding is noted not to stress differences in performance between different groups of children but to underline the importance of an assessment of this type of receptive language ability and the need for appropriate intervention procedures to reduce the risk of long-ranging disadvantage.

This point is further emphasized by the results of the four validation groups, where 31 per cent (N = 150), 26 per cent (N = 150), 18 per cent (N = 200), and 8 per cent (N = 100) of children completed all items of the Recognition of Initial Consonant Sounds Test correctly at the first administration of the test.

Table 5.2 shows how difficult each item of the Recognition of Initial Consonant Sounds Test was found by children in the different samples. The facility of each item is expressed in terms of the percentage of children in the sample who completed the item correctly. The arrangement of items for the final version of the test was determined by the order of difficulty experienced by the total representative sample. The item difficulties for the validation groups are presented in the last four columns, headed A, B, C, D, in Table 5.2.
Table 5.3 Recognition of Initial Consonant Sounds Test: Rasch-scaled Item Difficulty Estimates

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>57.4</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>53.2</td>
<td>1.1</td>
</tr>
<tr>
<td>3</td>
<td>53.0</td>
<td>1.1</td>
</tr>
<tr>
<td>4</td>
<td>52.3</td>
<td>1.1</td>
</tr>
<tr>
<td>5</td>
<td>51.7</td>
<td>1.1</td>
</tr>
<tr>
<td>6</td>
<td>50.5</td>
<td>1.2</td>
</tr>
<tr>
<td>7</td>
<td>50.0</td>
<td>1.2</td>
</tr>
<tr>
<td>8</td>
<td>49.8</td>
<td>1.2</td>
</tr>
<tr>
<td>9</td>
<td>49.2</td>
<td>1.2</td>
</tr>
<tr>
<td>10</td>
<td>49.0</td>
<td>1.2</td>
</tr>
<tr>
<td>11</td>
<td>49.0</td>
<td>1.2</td>
</tr>
<tr>
<td>12</td>
<td>48.5</td>
<td>1.2</td>
</tr>
<tr>
<td>13</td>
<td>48.3</td>
<td>1.2</td>
</tr>
<tr>
<td>14</td>
<td>48.1</td>
<td>1.2</td>
</tr>
<tr>
<td>15</td>
<td>47.8</td>
<td>1.2</td>
</tr>
<tr>
<td>16</td>
<td>46.1</td>
<td>1.3</td>
</tr>
<tr>
<td>17</td>
<td>46.1</td>
<td>1.3</td>
</tr>
</tbody>
</table>

a The differences in N observed between Tables 5.3 and 5.4 versus Tables 5.1 and 5.2 are because Rasch analyses exclude individuals who have obtained full scores or zero scores on the test.
Table 5.2 shows that a large percentage of children experienced difficulty with the task of recognizing initial consonant sounds, a task which is recognized as an important component of reading, spelling, and other language skills. As might be expected, native speakers of English do better on the task, but the skill can by no means be regarded as adequately established, even among school beginners whose language background is English.

Interpretation of Scores
Table 5.3 shows the estimates of item difficulty (Di) and the standard errors (SEDi) for the difficulties for all five samples as Rasch-scaled values. No significant differences were found between the Rasch item difficulties between the five groups (F = 0.02, df 4, 80, p > .05).

Inspection of Table 5.4 shows that the ability estimates obtained for each possible raw score on the test were extremely similar for the five separate and completely independent samples. Table 5.4 shows the stability of the Rasch-scaled ability estimates which can be obtained from the conversion of raw scores on the Recognition of Initial Consonant Sounds Test. The Rasch abilities (Ai) as obtained on the test by the five separate and independent samples of school beginners are shown. The standard error (SEAi) of the ability estimate is provided for each raw score level.

The correlations between the item difficulties and between the ability estimates based on raw scores were very high between the five independent samples. On the left-hand side of Table 5.5, the correlations between the Rasch item difficulties obtained in the five samples are shown; on the right-hand side of the same table the correlations between the Rasch ability estimates for the five groups are shown. Inspection of Table 5.5 shows the correlations between the item difficulties between groups to vary, but to be moderately high. The reason for their variation lies in the between-group variation of item difficulty. As the summaries of conventional item analysis procedures provided in Tables 5.1 and 5.2 showed, the performance levels of the groups differed.

The Rasch ability estimates obtained in the different groups correlated perfectly, as reflected by the coefficients of $r = 1.00$ presented in the right-hand side of Table 5.5. As noted previously, this finding emphasizes the stability of measurement, both in terms of its validity and its reliability. These results show that the Rasch-scaled ability estimate provided by the test will be constant for each individual, no matter in what type of group he or she is tested. For a more detailed discussion of points raised in this section, and of the assumptions and advantages of Rasch scaling, refer to Chapter 10 of this Handbook.

Implications
As with the abilities assessed in the Auditory Discrimination Test, the skills covered by the Recognition of Initial Consonant Sounds Test are considered to be developmental in nature. The case for educational intervention was argued in Chapter 4 and will not be repeated here. In brief, the basic need for early intervention presents itself because the child who, for whatever temporary or more
Table 5.4  Recognition of Initial Consonant Sounds Test: Rasch Ability Estimates (\(\hat{A}_i\)) and Standard Errors (\(SE_{\hat{A}_i}\)) for Five Independent Samples

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Representative sample</th>
<th>Validation samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{A}_i)</td>
<td>(SE_{\hat{A}_i})</td>
</tr>
<tr>
<td>1</td>
<td>36.6</td>
<td>5.0</td>
</tr>
<tr>
<td>2</td>
<td>40.2</td>
<td>3.6</td>
</tr>
<tr>
<td>3</td>
<td>42.5</td>
<td>3.1</td>
</tr>
<tr>
<td>4</td>
<td>44.4</td>
<td>2.8</td>
</tr>
<tr>
<td>5</td>
<td>45.8</td>
<td>2.6</td>
</tr>
<tr>
<td>6</td>
<td>47.1</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>48.3</td>
<td>2.4</td>
</tr>
<tr>
<td>8</td>
<td>49.4</td>
<td>2.4</td>
</tr>
<tr>
<td>9</td>
<td>50.6</td>
<td>2.4</td>
</tr>
<tr>
<td>10</td>
<td>51.7</td>
<td>2.4</td>
</tr>
<tr>
<td>11</td>
<td>52.9</td>
<td>2.5</td>
</tr>
<tr>
<td>12</td>
<td>54.2</td>
<td>2.6</td>
</tr>
<tr>
<td>13</td>
<td>55.7</td>
<td>2.8</td>
</tr>
<tr>
<td>14</td>
<td>57.5</td>
<td>3.1</td>
</tr>
<tr>
<td>15</td>
<td>59.8</td>
<td>3.6</td>
</tr>
<tr>
<td>16</td>
<td>63.4</td>
<td>5.0</td>
</tr>
<tr>
<td>N</td>
<td>132</td>
<td>132</td>
</tr>
</tbody>
</table>
Table 5.5 Recognition of Initial Consonant Sounds Test: Correlations between Item Difficulties and between Ability Estimates Obtained for Five Independent Samples of School Beginners

<table>
<thead>
<tr>
<th>Sample</th>
<th>Item difficulty (Di)</th>
<th>Ability estimate (Ai)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Representative</td>
<td>.33</td>
<td>.59</td>
</tr>
<tr>
<td>Validity A</td>
<td>.69</td>
<td>.72</td>
</tr>
<tr>
<td>Validity B</td>
<td>.87</td>
<td>.92</td>
</tr>
<tr>
<td>Validity C</td>
<td>.89</td>
<td></td>
</tr>
</tbody>
</table>

permanent reason, has greater difficulty in hearing and understanding the spoken word than his or her peers, is likely to be disadvantaged right from the beginning of his or her school career.

The user of either of the auditory analysis skills tests included in the ACER Early School Series is advised to study the total material in all sections of this Handbook which pertain to either of these tests.

Training in auditory reception and analysis skills needs to include practice in listening and the discrimination and reproduction of speech sounds, words, and sentences.

The following types of specific exercises might be found useful:

1. identification of specific sounds (e.g. animal sounds, car/motorbike sounds, household sounds, toy sounds, music, etc.);
2. discrimination of one sound from a background of others (e.g. whistle against traffic, solo instrument and orchestra, etc.);
3. listening for meaningful sounds in the environment;
4. following of verbal instructions (starting off with one task, directions presented in short sentences, attempt to utilize pictures and other visual clues);
5. learning of rhymes and rhythms.

However, the consideration of one important diagnostic possibility must not be neglected. It has to be remembered that children with auditory receptive difficulties which do not appear to improve with initial training are more likely to be suffering from hearing defects than individuals in the general population. Some such defects can be temporary, for example resulting from ear infections, colds, etc. Others may well respond to long-term medical treatment or be compensated by mechanical devices. Early referral for medical assessment would thus seem essential for suspected deficits of this type.
Area B: Conceptual Skills

Concrete and abstract concepts constitute one of the essential aids, both in the organization of human experiences and in higher cognitive functioning.

The importance of the development of conceptual skills in children has been recognized for some time (Lovell, 1962; Sigel, 1964; Vinacke, 1951). Russell (1956) pointed out:

The clarity and completeness of a child's concept are the best measure of his possible success in school learning because meaning is fundamental to such learning. (p.120)

Piaget (1963) identified various types of concepts which he found to be attained by children at different stages of development (Flavell, 1963). In her review of research concerned with the development of spatial concepts and classification, Maccoby (1964) noted considerable support for Piaget's theory of age-related qualitative changes in the conceptual development in children. She also noted the possible influence of individual and cultural differences on the rate of concept attainment (e.g. Braine, 1959; Duckworth, 1964; Wallach, 1963), a finding which has been emphasized increasingly in more recent times (for reviews see Goodwin and Driscoll, 1980; Sattler, 1974).

While Piaget's theory of stages in cognitive development has received considerable support in the research literature (e.g. Dodwell, 1960; Elkind, 1961; Flavell, 1963; Laurendeau and Pinard, 1962) a number of studies have failed to support Piaget's stages of development (Estes, 1956).

A number of other variables, apart from age, have been related to the development of concepts in children. These variables include previous experience, cultural factors, motivation, and training (Inhelder, 1962) and socio-economic background (Anastasi, 1958; Ausubel, 1958; Estes, 1956; John, 1963; Siller, 1957; Wilson, 1963).

The ACER Early School Series tests discussed in the following two chapters are the Number Test in Chapter 6 and the Figure Formation Test in Chapter 7. Both tests require the application of conceptual skills to non-verbal stimuli. The testee must reason to complete the items of these tests successfully.

The development of perceptual organization abilities (i.e. analysis and synthesis
skills such as the grouping of objects, the understanding of quantity, size, shape, space, and functional relationships between the parts of a whole) plays a major role in the individual's performance on these tests. Other variables contributing to test performance include concentration, the ability to distinguish between essential and non-essential parts, and the ability to use step-by-step sequential strategies.
THE NUMBER TEST

Rationale
The research literature contains a considerable amount of evidence for the view that the number scheme is a central quantity scheme which influences the way in which children form concepts in other areas (Elkind and Flavell, 1969; Gal'perin and Georgiev, 1969; Inhelder and Sinclair, 1969; Laurendeau and Pinard, 1970; Piaget, 1952, 1968; Wohlwill, 1962). These investigators tend to suggest 'that the processes which enable the child to apply the number scheme in general are essentially those which render quantities countable' (Gelman, 1972, p.164).

The hypothesis that counting is fundamental not only in arithmetic but also as a basic mechanism used by young children to estimate numbers of all sizes appears to be well supported (Beckwith and Restle, 1966; Gelman, 1972a, 1972b; Neisser, 1966; Smedslund, 1966). The term 'counting' tends to be used for two quite separate processes: the activity of repeating numbers by rote, and the ability to coordinate successive items or events with numerals (i.e. to represent the quantity of a given set of items by a numerical value). The latter is the meaning intended wherever the term is used in this Handbook.

The Number Test in the ACER Early School Series is aimed to assess the individual's ability to count objects. This activity 'requires a perceptual grouping of objects into those already counted and those still ahead' (Beckwith and Restle, 1966, p.437), and the attaching of correct numerals to express quantity. Items included in the test require both the counting of discrete objects and the counting of sets of objects. The ability to perceive assemblies of discrete objects in groups and the function of such groups in counting has long been described as an important perceptual (Wertheimer, 1945) and developmental stage (Woodworth and Schlosberg, 1955). Its application transcends perceptual development. The ability to process discrete objects as sets is of particular relevance in the development of the estimation of quantity without counting, in the development of classification skills, and in concept formation more generally (Piaget, 1952; Siegel, 1971; Wohlwill, 1960).
THE NUMBER TEST

Description
The Number Test is an eight-item pictorial test. Each item requires the child to represent the quantity of discrete objects, or sets presented by drawing an equivalent number of 'sticks' into a 7.5 x 7.5 cm 'box' provided for this purpose. The items to be counted consist of simple everyday objects (e.g. cars, planes, trucks, houses, dogs, etc.). The drawings are simple and relatively large. Two items are represented on each page of the disposable test booklet.

Development
An initial item pool of 20 items was tried out with individuals and groups of three to 20 preparatory grade and kindergarten children in three schools and two kindergartens. A number of teachers volunteered to try out the test with school beginners of their choice. More than half of the items appeared to be suitable for inclusion in the test. The final selection of items was based on the preferences expressed by the teachers and kindergarten teachers who had co-operated in the trials and on the advice received from school counsellors and guidance officers concerned with school beginners in the Australian Capital Territory, New South Wales, Victoria, and Queensland, who emphasized the need for the tests in this battery to be as short as possible.

Various types of directions and a variety of wordings were tried out during initial trials of the test.

Administration and Scoring
The administration of the practice and test items is preceded by three preparatory exercises which are intended to convey to the child both the activity of counting and the representation of the counted objects by means of stick drawings. The Directions for Administration contain full instructions for these preliminary exercises. The child marks his or her responses to these exercises on the reverse side of the test booklet which contains three boxes of decreasing size, the smallest size (7.5 x 7.5 cm) being equivalent to the size of the boxes accompanying the test items.

The time required for administration of the test varies with the maturity and/or handicap of the testee and the size of the group of children tested at one time. During trial testing the average total time required to administer the Number Test to groups of three to eight 5- or 5½-year-old school beginners was 10 to 15 minutes, excluding the preliminary exercises. The time required for the latter varied from 2 to 5 minutes.

Like the other tests in the ACER Early School Series, the Number Test is a power test, not a speed test. Trial testing showed, however, that the performance of very slow children did not improve when additional time was given. It is suggested, therefore, that except in cases of particular physical handicap, no more than 20 minutes be allowed for the completion of the six test items.

Materials required consist of the Directions for Administration which contain general and detailed instructions for test administration, a test booklet, cardboard marker, crayon or pencil for each child, and a spare test booklet which the teacher can use for demonstration purposes. The instructions are in simple English and are printed out in full for each item.
Table 6.1 Number Test: Performance of 192 School Beginners

<table>
<thead>
<tr>
<th>Mode</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 and 8 (58%)</td>
<td>7 (30%)</td>
<td>8 (33%)</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Mean</td>
<td>6.3</td>
<td>5.6</td>
<td>6.5</td>
</tr>
<tr>
<td>SD</td>
<td>1.7</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>79.1%</td>
<td>67.7%</td>
<td>81.4%</td>
</tr>
<tr>
<td>SD</td>
<td>14.2%</td>
<td>21.8%</td>
<td>12.4%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.70</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>N</td>
<td>192</td>
<td>156</td>
<td>36</td>
</tr>
</tbody>
</table>

Scoring

A score key is provided. One point is given for each correct item. Possible points are zero (0) or one (1) for each item. The raw score for the test is provided by the number of correct items. The maximum obtainable raw score on the Number Test is 6. Tables 6.4 and A.3 allow the conversion of raw scores to Rasch-scaled developmental ability scores.

Research Results

Table 6.1 presents a description of the performance of 192 children of the representative sample of 197 school beginners described in Chapter 3.

In the total sample, 29 per cent of children completed all items correctly (score of 8), and another 29 per cent obtained a score of 7. The difference in mean performances between children from non-English and English language backgrounds was statistically significant ($t = 2.91, df = 190$). The reason for this may be simply that some of the non-English background children were unable to follow the directions because of language difficulties. On the other hand, Table 6.2 shows that the non-English background group found the first three items no more difficult than did the English background group, and that the discrepancy in the performances of the two groups increased with increasing difficulty of the items. This suggests that it can be assumed that the non-English group had understood what was required by the tasks, and might lead to hypotheses of possible developmental disadvantage in the non-English group. As the proportion of non-English background children was low in the present sample, these statistics must not be over-interpreted or generalized. The intention of the above discussion is solely to demonstrate and emphasize the importance of the gathering of qualitative as well as quantitative information in diagnostic assessment. This matter will be discussed further under Implications later in this chapter.

A further finding which may be of general importance to teachers is that the greater average item difficulty for non-English background children was accompanied by a considerably larger standard deviation than was obtained in the
Table 6.2 Number Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Non-English</td>
</tr>
<tr>
<td>1 planes</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>2 houses</td>
<td>90</td>
<td>89</td>
</tr>
<tr>
<td>3 money</td>
<td>86</td>
<td>83</td>
</tr>
<tr>
<td>4 trucks</td>
<td>84</td>
<td>69</td>
</tr>
<tr>
<td>5 cars</td>
<td>81</td>
<td>69</td>
</tr>
<tr>
<td>6 men</td>
<td>77</td>
<td>67</td>
</tr>
<tr>
<td>7 shoes</td>
<td>76</td>
<td>67</td>
</tr>
<tr>
<td>8 ladies</td>
<td>47</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 6.2 shows the difficulty found by children in different samples for each item of the Number Test. The facility of each item is expressed by the percentage of children in the sample who completed the item correctly. The ordering of items for the final version of the test was determined by the performance of the total representative sample.

The column headed 'Validity sample' contains the results of a further 612 school beginners which were provided by 13 teachers and school counsellors, and had been obtained from unselected groups of school beginners in city and rural schools of varying size and socio-economic status. The proportion of non-English background children in these samples varied but was considerably higher than in the representative sample. The sample is discussed in Chapter 3. Initial separate analyses of the Number Test results of these 13 additional samples showed only minimal discrepancies in the rank ordering of items. One set of item difficulties was thus determined for the total validation sample.

**Interpretation of Scores**

Table 6.3 shows the estimates of Rasch item difficulty (Di) and standard errors (SED) of the difficulties for both the representative and the validation samples.

No significant difference was found in the mean Rasch item difficulty between the two groups (t = 0.28, df, p > 0.05). The correlation between the item difficulties for the two groups was extremely high (r = 0.98). This finding and the very close agreement of Rasch ability estimates (Ai) obtained for these two different and completely independent samples, as presented in Table 6.4 shows the stability of the Number Test as an assessment device for different samples. The
Table 6.3 Number Test: Rasch-scaled Item Difficulty Estimates

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SE(Di)</td>
</tr>
<tr>
<td>1</td>
<td>61.5</td>
<td>1.3</td>
</tr>
<tr>
<td>2</td>
<td>54.5</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>52.0</td>
<td>1.3</td>
</tr>
<tr>
<td>4</td>
<td>50.3</td>
<td>1.3</td>
</tr>
<tr>
<td>5</td>
<td>48.0</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>46.7</td>
<td>1.5</td>
</tr>
<tr>
<td>7</td>
<td>44.6</td>
<td>1.7</td>
</tr>
<tr>
<td>8</td>
<td>41.0</td>
<td>2.0</td>
</tr>
<tr>
<td>L</td>
<td>49.8</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>133</td>
<td></td>
</tr>
</tbody>
</table>

- The difference in N between Tables 6.3 and 6.4 and Tables 6.1 and 6.2 is because, in Rasch analyses, individuals who obtained full scores on the test are eliminated from the analysis.

The correlation between the ability estimates in the two groups was found to be perfect (r = 1.0).

Table 6.4 describes the performances of the two samples in terms of Rasch counting ability (Ai). An indication of the standard error (SEAi) is provided for the Rasch ability estimate of each raw score.

For a more detailed discussion of the Rasch model of measurement, its assumptions and advantages, and the fit of the ACER Early School Series tests to the model, refer to Chapter 10.

**Implications**

The results of the *Number Test* provide the teacher with a broad indication of the degree to which individuals and the group as a whole can count correctly. Observation of the procedure by which each child obtains the answer to each item is of further diagnostic importance, as it might provide the teacher with an indication as to the size of the number sets to which the initial teaching of arithmetical reasoning might need to be restricted in the case of particular individuals. Some children appear to have mastered the principles of counting for small sets of objects. If their ability to count correctly (i.e. their ability to represent accurately the size of a set a as a numerical value) is limited to small sets, they must be expected to have difficulty in keeping track of and manipulating larger set sizes when required to reason about them.

Further diagnostic information can be obtained by noting whether the child completes items by counting or by using one-to-one correspondence. The latter type of strategy is regarded as an indication of emerging counting skills, more frequently found in three- to four-year-olds than in five-year-olds (Gelman and Gallistel, 1978). Older children are more likely to use one-to-one correspondence strategies with larger than with smaller sets.
Table 6.4: Number Test: Rasch Ability Estimates

<table>
<thead>
<tr>
<th>Obtained raw score</th>
<th>Representative sample</th>
<th>Validation sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SE(Ai)</td>
</tr>
<tr>
<td>1</td>
<td>37.8</td>
<td>6.5</td>
</tr>
<tr>
<td>2</td>
<td>43.3</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>46.7</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>50.0</td>
<td>4.3</td>
</tr>
<tr>
<td>5</td>
<td>53.0</td>
<td>4.5</td>
</tr>
<tr>
<td>6</td>
<td>56.7</td>
<td>5.0</td>
</tr>
<tr>
<td>7</td>
<td>61.7</td>
<td>6.5</td>
</tr>
<tr>
<td>N</td>
<td>133</td>
<td>406</td>
</tr>
</tbody>
</table>

Sources of error vary, and they provide valuable information for the design of intervention procedures. In the trial testing studies, most errors were due to a tendency on the part of some individuals to double count or to skip an object in the set. These errors occurred among both children who counted consecutive objects directly, and those using one-to-one correspondence. As a matter of fact, errors of this kind among children using the latter type of strategy appeared frequently to be the result of some difficulty in co-ordinating the complex motor activity involved in this type of strategy of counting. Less mature children and individuals with delayed motor development or poor visual motor co-ordination had some difficulty in this activity which requires them to point and draw a line at the same time. Some children showed particular difficulties at the beginning and/or the end of certain items.

Gelman and Gallistel (1978) and Vygotsky (1978) have suggested that the type of object to be counted may not be irrelevant to the application of counting procedures in three- and four-year-old children. Results of item analyses of the present test would lend support to this hypothesis. Item 3, relating to piles of money, was found to be an easier item in all samples than the items which required the counting of larger numbers of discrete items. Reasons for this may be related to familiarity and perceived importance, as well as to the development of abstraction.

Observed differences in strategies used to complete different items in the test by one individual, or a refusal to complete particular items, especially in relation to performance on discrete versus set items, can be expected to provide the teacher with valuable guidelines in the planning of teaching programs. Other systematic sources of error might be observed in individuals or groups. All of these suggest that particular children require further practice in counting, even though they may have grasped the general principle of what is required. The goal is that each child’s understanding of quantity will generalize to a wide range of settings.

Extensive use of manipulative and concrete objects is recommended for initial counting practice at this level. Sequence and position may need to be taught. Dot-to-dot drawings and similar activities can be used to teach sequence. The abacus and various counting frames provide useful devices which children can use effectively.
FIGURE FORMATION-TEST

Rationale
The aim of the Figure Formation Test differs from the general purpose of the other tests included in the ACER Early School Series. In the other nine tests, prerequisite and component skills of basic learning processes are assessed. The performance of individuals in these tests is expected to provide the teacher with diagnostic information concerning the individual's strengths and weaknesses in the particular skills measured by each test. This information may then be used in the planning of learning activities and other educational experiences which meet the needs of the individual and cater for the identified individual differences of the group to be taught.

The contribution of the Figure Formation Test to the ACER Early School Series is to provide a quick but rough indication of the child's more general intellectual maturity. The latter term, as used here, is not intended to convey any notions of intelligence as a unitary construct on the basis of which individuals might be categorized or labelled. Rather the term intellectual as used here might be exchanged by the word 'conceptual', as proposed by Harris (1963):

By intellectual maturity is meant the ability to form concepts of increasingly abstract character. Intellectual activity required: (1) the ability to perceive, i.e. to discriminate likenesses and differences; (2) the ability to abstract, i.e. to classify objects according to such likeness and differences; and (3) the ability to generalize, i.e. to assign an object newly experienced to the correct class, according to discriminated features, properties, or attributes. These three functions, taken together, comprise the process of concept formation. (p.5)

A concept is usually defined in psychology as the product of a mental or thought process whereby the qualities, aspects, and relations of objects are identified, compared, abstracted, and generalized. When the process covers different individuals or items, it is called a class concept. When the process represents a common aspect of the class, it is an abstract idea. The processes which include perception, conceptualizing, and knowing, as well as judging and reasoning, are called cognition. (p.6)

More mature and older children (although age and experience do not appear to be the only factors) are able to discern and specify the properties of objects and situations in greater detail. They are also able to recognize the characteristic and essen-
tial properties and relationships of objects, classes of objects, and eventually abstract ideas more readily, and tend thus to be able to form more precise and generalizable concepts.

The inclusion of a test of more general conceptual ability in the ACER Early School Series was requested particularly by school counsellors and infant teachers. Traditionally those wishing to include in the assessment of school beginners a measure of more generalized conceptual maturity have tended to administer the Goodenough-Harris 'Draw-a-Man' test (Goodenough, 1926; Harris, 1963).

In recent years an increasing amount of disenchantment with this test has been expressed both in the literature (e.g. Ernhart, 1975; Goodwin and Driscoll, 1980) and by practitioners. Criticisms have mainly been related to the complicated requirements for the scoring and interpretation of the test, the negative influence of cultural and other experiential factors on the scores obtainable on the test, and questions relating to a possible decrease in the validity of the test in recent years, during which imagination and creativity have tended to be emphasized in children's drawings and language skills.

The Figure Formation Test was designed to provide a quick evaluation of the child's understanding of basic concepts and relationships. The ability to identify necessary parts belonging to a familiar object will reveal the discrimination the child has made about that object and its parts as belonging to a category, in other words, as a concept.

Specific skills required for the completion of the tasks making up the Figure Formation Test include the ability to visualize separated parts as a whole unit. The child needs to be aware of the relationships between space, shape, size, and the figure as a whole. Another ability involved in the successful performance on the Figure Formation Test is the ability to relate three-dimensional objects of the real world to two-dimensional stimuli as presented in the items of the test.

Where the test is administered individually or to small groups, observation during the child's work on each item can provide the teacher with diagnostic information concerning the child's conception of the particular item, and of the appropriateness of the strategies used in completing the item. Inability to recognize the figure and trial-and-error responses might be noted. The types of errors most frequently made by an individual throughout the test provide useful guidelines to the design of educational programs.

The opportunity for such close observations of the performances of individuals will probably occur rarely in the screening of school beginners. The assumption is that the test will be administered to reasonably large groups. The simplicity of the instructions and the tasks themselves, together with the brevity of this test, make the Figure Formation Test suitable to be administered to larger groups than is advised in the case of the other tests contained in the ACER Early School Series.

It is suggested, therefore, that the diagnostic component of the Figure Formation Test as used in the ACER Early School Series (for other uses, see the section on Development) will tend to originate less from variables inherent in the test performance itself, but rather from the consideration of the score on the Figure Formation Test in combination with the child's performance on other tests, particularly the
THE TESTS

Number Test. The implications of the performance on the Figure Formation Test and its diagnostic use are discussed further at the end of this chapter.

Description

The Figure Formation Test is a four-item pictorial test which can be administered to individuals or to a group. Throughout the test the individual is required to identify shapes and the functional relationship between them by marking the essential component parts of a familiar object. In keeping with the school beginners' level of development, the items selected are concrete rather than abstract at this stage.

Each item consists of a single page, the upper half of which shows the complete object, and the lower half of which contains parts which make up the object and a number of spare parts. The testee is required to mark the parts which are necessary to make up the object. The drawings are large and simple.

In both content and administration, the Figure Formation Test departs from the style of the other tests contained in the ACER Early School Series. As previously noted, the aim of the test is to assess broader conceptual maturity rather than abilities which are specifically related to a particular area of the child's learning at school. The administration of the test is extremely simple and, unlike the other tests, the Figure Formation Test can be administered to large groups of children at one and the same time.

Development

Different types of items and assessment procedures were considered. In the initial stages of development, the methodology used was that of asking children to draw objects. The difficulties in scoring and interpreting these responses led to the adoption of the method of asking the child to identify the relevant parts of objects from a given set.

Special consideration was given to different methods of scoring the responses. A number of different procedures were tried out. Finally, a decision had to be made between two different methods of scoring. It was found that a rapid evaluation of the child's understanding of basic concepts and the relationships between parts of an object can be made by scoring each item as either correct or incorrect.

Another method, which is more precise but also more time-consuming and demanding as far as the tester is concerned, would have been to award point scales for each item. Proportional scores would, in this case, be awarded for partly correct responses. Points would be deducted for wrongly marked parts, etc. In other words this latter scoring procedure would be similar to that of tests like the Goodenough-Harris 'Drawing Tests' (Goodenough, 1926; Harris, 1963). Such a method of scoring was not completely dismissed as a suitable procedure for the scoring of the Figure Formation Test.

However, for the use made of this test in the ACER Early School Series, where the aim was to provide a set of reasonably straightforward, simply scored and interpreted screening and diagnostic tests, it was decided to adhere to dichotomous scoring of each item, as applied to the other nine tests contained in the series. As part of the screening and diagnostic procedures provided by the ACER Early
School Series the results obtained by means of the *Figure Formation Test* will provide the counsellor and teacher with a quick assessment of the child’s ability to perform the tasks of which the test is comprised. The results of this test will also provide a rapid means to identify those children who have more difficulty with the tasks than their peers. The performance of the latter group should be further investigated for diagnostic signs, so that appropriate teaching and compensatory procedures can be initiated.

If the *Figure Formation Test* were to be used not solely to provide a rough indication of the conceptual maturity of individuals, but as a developmental test identifying different levels of intellectual maturity, the use of a scoring procedure of the second type discussed above would be more appropriate. Further work in this direction is in progress at present. Initial indications are that the *Figure Formation Test* may well form part of another series of ACER tests which are being prepared, namely a set of non-verbal measures aimed to assess general ability at various stages of development.

The test in its present form was tried out with a representative sample of 193 school beginners. The characteristics of this sample were described in Chapter 3. The results of an additional 600 school beginners on the *Figure Formation Test* were obtained in the form of four samples of unselected subjects. These data were provided by a number of teachers from unselected government and Catholic parish primary schools. The proportion of children from non-English language backgrounds could not be ascertained in these data.

**Administering and Scoring**

Administration of the *Figure Formation Test* is extremely simple. For example, for the practice item the tester says: ‘Now put a cross on all the parts which are needed to make up the boat’.

Unlike the other tests in the ACER Early School Series, this test can be given to quite large groups at the same time. During the trial testing, the test was administered successfully to classes of 15 to 20 school beginners early in their first term at school.

The administration of the *Figure Formation Test* does not require an experienced tester. The person administering the test need only have experience with children and be competent at following directions. The Directions for Administration contain full instructions. A teacher’s aide or other assistant could be asked to administer this test.

Materials required consist of the Directions for Administration, which contain general and detailed instructions for test administration, a test booklet, and a crayon or pencil for each child, and a spare test booklet which the tester can use for demonstration purposes.

The time required for administration of the total test varies with the maturity and experience of the testees and the size of the group tested at one time. During trial testing, the average time required for this test was five minutes.

The *Figure Formation Test* is a power test, not a speed test. However, it has been found that the performances of very slow-working children did not improve when
THE TESTS

Table 7.1 Figure Formation Test: Performance of a Representative Sample of 193 School Beginners

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total sample</td>
</tr>
<tr>
<td>Mode</td>
<td>4 (48%)</td>
</tr>
<tr>
<td>Median</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td>3.0</td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>77.2%</td>
</tr>
<tr>
<td>SD</td>
<td>4.8%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.65</td>
</tr>
<tr>
<td>N</td>
<td>193</td>
</tr>
</tbody>
</table>

extra time (beyond six to seven minutes for the total test) was allowed. Therefore it is suggested that, except in cases of special handicap, the total test be administered in approximately five minutes.

Scoring

A score key is provided. One point is given for each correct item. Possible points are zero (0) or one (1) for each item. The raw score for the test is provided by the number of correct items. The maximum score for the test is 4. Tables 7.4 and A.3 allow the conversion of raw scores to Rasch-scaled developmental ability scores.

Research Results

In this section, a summary is provided of the results obtained on the Figure Formation Test by the representative sample of 193 school beginners who were tested five weeks after the beginning of their preparatory school year, and of four unselected validation groups, previously described in the section on Development.

Table 7.1 summarizes the performance of the representative sample of school beginners, and shows that 54 per cent of the children whose home language was English completed all items of the test correctly. Only 24 per cent of the children from non-English language backgrounds completed all items correctly.

Among the four comparison samples, 74 per cent (N = 50), 57.5 per cent (N = 400), 45 per cent (N = 100), and 16 per cent (N = 50) of samples A, B, C, and D respectively completed all items on the test correctly.

Table 7.2 shows how difficult each item of the Figure Formation Test was found
FIGURE FORMATION TEST

Table 7.2 Figure Formation Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Non-English</td>
</tr>
<tr>
<td>1 truck</td>
<td>82</td>
<td>66</td>
</tr>
<tr>
<td>2 teapot</td>
<td>80</td>
<td>66</td>
</tr>
<tr>
<td>3 robot</td>
<td>76</td>
<td>53</td>
</tr>
<tr>
<td>4 church</td>
<td>71</td>
<td>53</td>
</tr>
<tr>
<td>Average item facility</td>
<td></td>
<td>77.2</td>
</tr>
<tr>
<td>SD</td>
<td>4.8</td>
<td>7.6</td>
</tr>
<tr>
<td>N</td>
<td>193</td>
<td>38</td>
</tr>
</tbody>
</table>

by children in the five different and completely independent samples. In this table, the facility of each item is represented in terms of the percentage of school beginners in the sample who completed the item correctly.

The arrangement of items for the final version of the test was determined by the order of difficulty experienced by the total representative sample.

Table 7.2 shows that a considerable number of school beginners experienced difficulty with all items, and that the samples varied in ability. The non-English language background children’s performance was considerably weaker than that of their peers whose home language was English. It is not surprising to find that language difficulty affects conceptual maturity. This phenomenon has been addressed in the research literature, and will be discussed in the next chapter.

Interpretation of Scores

Table 7.3 contains the estimates of item difficulty ($D_i$) and the standard errors ($SE_{Di}$) for the item difficulties for all five samples as Rasch-scaled values. No significant differences were found in the Rasch-scaled difficulties of the items for the five groups.

Table 7.4 shows the ability estimates ($A_i$) obtained for each possible raw score on the Figure Formation Test, and the standard error ($SE_{Ai}$) for each ability estimate.

The ability estimates for the different samples show a very high stability. This shows that Rasch scales allow the estimation of ability independent of the sample as part of which an individual is assessed. While some variation between groups can be expected as regards the item difficulties — because of ability differences between the groups — the ability estimates should be, and were found to be, stable across groups. The latter finding is also reflected in Table 7.5, which shows the correlations between Rasch item difficulties obtained in the five samples on the left-hand side of the table, and the correlations for the ability estimates for each raw score between the five groups on the right-hand side of the table.

Inspection of Table 7.5 shows that the correlations between the item difficulties are moderately-high to high. As noted previously, the reason for the variation of the item difficulty correlations lies in the between-group variation of item difficulty.
Table 7.3  Figure Formation Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SEDi) for Five Independent Samples

<table>
<thead>
<tr>
<th>Item</th>
<th>Representative sample</th>
<th>Validation samples</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SEDi</td>
<td>Di</td>
<td>SEDi</td>
</tr>
<tr>
<td>1</td>
<td>52.6</td>
<td>1.1</td>
<td>55.5</td>
<td>2.7</td>
</tr>
<tr>
<td>2</td>
<td>50.7</td>
<td>1.1</td>
<td>55.5</td>
<td>2.7</td>
</tr>
<tr>
<td>3</td>
<td>48.8</td>
<td>1.1</td>
<td>46.4</td>
<td>3.8</td>
</tr>
<tr>
<td>4</td>
<td>47.9</td>
<td>1.1</td>
<td>42.6</td>
<td>5.1</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td></td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>2.1</td>
<td></td>
<td>6.5</td>
<td></td>
</tr>
</tbody>
</table>

Note: The differences in N observed between Tables 7.3 and 7.4 versus Tables 7.1 and 7.2 are because Rasch analyses exclude individuals who have obtained full scores or zero scores on the test.

Table 7.4  Figure Formation Test: Rasch Ability Estimates (Ai) and Standard Errors (SEAi) for Five Independent Samples

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Representative sample</th>
<th>Validation samples</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SEAi</td>
<td>Ai</td>
<td>SEAi</td>
</tr>
<tr>
<td>1</td>
<td>44.8</td>
<td>5.4</td>
<td>43.4</td>
<td>6.9</td>
</tr>
<tr>
<td>2</td>
<td>50.0</td>
<td>4.7</td>
<td>50.0</td>
<td>6.0</td>
</tr>
<tr>
<td>3</td>
<td>55.2</td>
<td>5.4</td>
<td>56.6</td>
<td>6.9</td>
</tr>
<tr>
<td>N</td>
<td>89</td>
<td>89</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>

Note: The differences in N observed between Tables 7.3 and 7.4 versus Tables 7.1 and 7.2 are because Rasch analyses exclude individuals who have obtained full scores or zero scores on the test.
for some items. As the results of classical item analysis procedures showed in Table 7.2, the performance levels of the groups differed.

The Rasch ability estimates obtained in the different groups correlated perfectly. As noted previously, this result emphasizes the stability of the Rasch-scaled measurement both in terms of its validity and its reliability. For a more detailed discussion of the latter point, refer to Chapters 3 and 10 of this Handbook.

Implications

As noted previously, the Figure Formation Test aims to provide a rough indication of the child's conceptual maturity. The test does not relate closely to any specific aspect of the school beginners' learning program. However, as is the case with most tests of more general intellectual functioning, the abilities required for successful completion of the items of the Figure Formation Test have undoubtedly some bearing on other areas of perception and learning.

The ability to recognize familiar objects and to separate essential from non-essential parts of the whole (Sattler, 1974) and, perhaps more importantly, the ability to visualize the component parts of a concrete object and to assemble these parts into an imagined whole, involve visual closure. In other words, the child must be able to visualize the relationship of fragmented parts into a whole without actually physically putting the parts together.

Techniques to aid the development of visual closure would emphasize familiar objects and the relationship of parts to the whole. Manipulation of objects by the child is important. Jigsaw puzzles of familiar objects, sewing cards, pictures which can be cut into pieces and reassembled, and a variety of games can help. Simple kits which can be built by the children, who may subsequently be asked to explain various parts of what has been built, and their function. Other exercises of visual closure might include the following types of activities:

1. the identification of parts of the body
2. the search for and identification of objects in pictures
3. the identification of items or parts of items which have been represented incorrectly in a picture
4. the completion of incomplete drawings
5. matching of objects, colours, shapes, symbols, geometric forms, letters, etc.
6. the identification of groups or classes of objects in pictures.

An understanding of position in space is another ability assessed in the Figure Formation Test. In terms of academic skills, the child who has difficulty in the
understanding of spatial relationships of the type presented in the items of the Figure Formation Test may be the individual who, without special attention from the teacher, might have difficulty in differentiating between similar letters, such as b and d; who might not recognize the proper sequence of letters in words (e.g. who might reverse letters and read 'saw' for 'was', or interchange letters in other ways); who might not read or write letters in proper sequence within a sentence; or whose letter formation in writing may be characterized by rotations and mirror writing.

Another ability which may influence the performance on the Figure Formation Test in some individuals is figure-background discrimination: that is, the ability to perceive a stimulus apart from its background. An individual who is deficient in this aspect of visual perception may perceive all components of a picture with equal strength. In other words, he or she will be unable to focus on a central theme or figure as opposed to the background.

Children who are lacking in figure-ground discrimination ability often appear disorganized and distracted. However, this type of distraction is not so much a matter of shifting attention, as the fact that the individual cannot distinguish background material from the stimulus under consideration, and thus loses the meaning of the stimulus and its relevance. Background material is moved into the foreground and weakens, displaces, or even masks the stimulus. This deficiency occurs in adults as well as in children, and can occur in other sensory areas (e.g. in auditory discrimination).

The consideration of performance scores obtained on the Figure Formation Test in relation to the performance on the other tests of the ACER Early School Series has been recommended earlier in this chapter.

Among children in the research samples, the test provided valid diagnostic information concerning various types of disadvantages. The finding of a moderately high to high score on the Figure Formation Test combined with very low scores in the language, or in the language and auditory analysis tests identified children who, with extra help, were able to increase their language skills very much faster than had been expected.

As a result of information obtained during trial testing, the performance of a group of children with very low scores on the Number Test, but adequate to high scores on the Figure Formation Test, was investigated further. It was shown that the low scores these individuals had obtained on the Number Test were due to technical and motor problems, rather than to an inability to perform the counting tasks required by the items of the Number Test (Turner, 1978).
Area C: Language Skills

Language skills are of central importance for human functioning. The facilitation of language development is thus a vital concern for parents and all educators. Although many divergent theories of language acquisition can be identified, a dominant view has been expressed by Ruder (1972):

It is apparent that a theory of grammar must include phonology, syntax and semantics. This is taking a rather broad view of grammar or of language, but it is upon such a general viewpoint of grammar that this psycholinguistic model of language acquisition is based. (p.5)

Other theorists have maintained that three levels of language can be identified—syntactic, a semantic, and a phonological level—and that, although each is governed by separate rules, all three levels are integrated (Roberts, 1964; Smith, 1973; Smith and Miller, 1966; Solberg, 1973).

Perkins (1971), citing Lenneberg (1969), elaborated upon the nature of language and upon the interrelationship of its four components—phonemic, morphemic, syntactic, and semantic elements and their changing interrelationships.

Not a static product of the mind, it is the dynamic manifestation of the functioning of his brain. Every level of language—phonemic, morphemic, syntactic, and semantic—facilitates the discrimination of relations among classes of objects as they belong to different classes. Never is a word a unique name for a single object. So in the broadest sense, a disorder of language impairs the ability to discriminate relations. (Perkins, 1971, p.238)

It has been emphasized frequently that children whose command of language is adequate or advanced seem to be able to cope better with the demands made by their environment, especially with school learning. Until relatively recently less attention appears to have been given to frequently replicated research findings which have shown the importance of language development among early environmental influences, in the development of thought processes and intelligence in young children (Carroll, 1966; Cazden, 1972; Hunt, 1961; Vygotsky, 1962). These researchers argued that delayed, inadequate, and restricted language development in young children has led not only to an
incomplete knowledge of syntax and to a severely restricted vocabulary in later life, but to lasting disadvantage (Bernstein, 1961; Labov, 1970).

For a large proportion of children in Australian schools, English is not the native language, nor the language spoken in the home. In many other cases the type or style of English used in the home does not match the English used in the classroom. The phenomenon, irrespective of whether it is caused by a language difference, a developmental deficit, or a combination of both, is characterized aptly by Cazden (1972):

The language problems of the educationally disadvantaged child are not merely problems of expression or dialect. They are basically conceptual problems... Most of these children will remain confused throughout their school career unless they are taught, starting now, the basic concepts and the manner in which these concepts are expressed through the kind of language used in school. Most of these children will be severely restricted in their future ability to think and generalize unless they are taught something about the basic structure of language and the basic concepts that will be used in the classroom from the first day they enter the first grade... Educational programs in this country are predicated on the student's ability to use... complex (standard English) grammar. It is the form of English used in conveying information to students and it is expected to be reproduced in their own writings and communication process. It is not just a matter of their having a different vocabulary due to different life experiences, but also of their not having the same flexibility in manipulating words grammatically (p.62)... The student is handicapped in his ability to communicate and to receive communications. Without an adequate grasp of language, it is impossible to cope with abstract ideas. For this reason, children must be encouraged to develop their language. It enables them to think. (p.63; italics not in original)

It has been shown that early intervention can foster and expand the language skills of disadvantaged and developmentally retarded children (Bereiter and Engelmann, 1966; Blank and Solomon, 1968; Dale, 1965; Drash, 1972; Eisenson, 1972; Guess, Sailor, and Baer, 1974; Klaus and Gray, 1968; Luria and Yudovich, 1971; Sloane and MacAuley, 1968). The basic aim of such intervention is to increase the child's ability to understand language in a wide variety of areas, contact with some of which the teacher may be providing for the child for the first time. An effective language development program must be based on the teacher's knowledge of language strengths and weaknesses of all children in the group.

In considering possible areas for the valid assessment of language skills in school beginners, an attempt was made to select abilities which could be regarded as basic to the child's ability to receive verbal communications and which, at the same time, could be regarded as readily trainable. Cazden (1971) following Chomsky (1965) stressed the importance of developing simple and more difficult language skills in syntax and vocabulary as a first step.

In an empirical pilot study, a number of language samples were collected from four-year-old native speakers of English who were attending kindergarten, and from children attending the same kindergarten whose home language was not English. Content analyses of the recorded samples resulted in the identification of vocabulary, knowledge of prepositions, pronouns, verb tense, and general comprehension as major areas in which significant-performance discrepancies between competent and advanced language users and linguistically disadvantaged children could be observed.

A careful review of the literature supported this emphasis and led to the
selection of the areas of syntax, comprehension, and word knowledge as domains for the development of the tests of language skills in the ACER Early School Series.

The specific syntactic structures to be assessed were selected so as to coincide with the types of structures which had tended to be regarded as fundamental in syntax development in this age group in the research literature (Berko, 1958; Brown, 1973; Brown and Fraser, 1964; Cazden; 1968; Fraser, Bellugi and Brown, 1963; Menyuk, 1963; Miller, 1973). Five separate tests were developed, each of which emphasized one of the following syntactic structures: Prepositions, Verb Tense, Pronouns, and Negation. These tests and the Comprehension Test are described and discussed in Chapter 8. Chapter 9 discusses the Word Knowledge Test.

The method of presentation adopted in Chapter 8 departs somewhat from that of the previous chapters, in that the five tests are discussed together. The structure of Chapter 8 is the same as that of Chapters 4 to 7, and of Chapter 9. The seven section headings — Rationale, Description, etc. — used in the discussion of the tests presented in Chapters 4 to 7 and in Chapter 9 have been adhered to in Chapter 8. However, instead of organizing the discussion of one test after another, Chapter 8 is structured on the basis of the section headings, and the relevant information concerning all five tests is reported together under each heading.

The reason for this departure from the general design is that the tests of syntactic structures are less independent conceptually than the other tests. They all assess competence in syntax although each test focuses on a different syntactic domain.

It is assumed that teachers wishing to assess language skills will usually make use of all five tests. This provides another reason why it may be appropriate to present the rationale, description, implications, etc. of these assessment instruments in one chapter.
TESTS OF SYNTAX

Rationale
The general rationale for the inclusion of tests of syntax in the ACER Early School Series was provided in the introductory pages on language skills in the previous section of this Handbook. The choice of the syntactic structures to be covered by these tests was determined by published research and on the basis of empirical pilot data obtained from four- to four-and-a-half-year-old children in Australia. The validity of the items constituting syntax tests is based on published studies which are generally regarded as having provided replicable and authentic descriptions of the acquisition of syntactic structures in the language of young children (Brown, 1973; Brown and Fraser, 1964; Cazden, 1968; Miller, 1973) and research which has investigated and described the control young children have over various aspects of English syntax (Berko, 1958; Menyuk, 1963, 1964).

A common assumption held until the late 1960s was that, by this stage, five-year-olds have mastered the syntactic structures of their native language, and that later developments consist mainly of the addition of a sophisticated vocabulary. This view was challenged by Piaget’s research, which stressed that language development is dependent on more general mechanisms governing the child’s overall cognitive development. His studies on various categories of knowledge showed that many crucial developments in language and cognition take place well beyond the age of five years. Sinclair’s (1967) study showed very clearly that many components of syntactical structures (e.g. prepositions, pronouns, verb tenses) though present in a young child’s vocabulary were not fully internalized nor used as relational terms until the child’s corresponding concepts in conservation and
seriation had been mastered. On the basis of psycholinguistic research, Carol Chomsky (1969) showed that the real understanding and mastery of complex syntactical structures, such as complex adjectival, verbal, and pronoun constructions, develop as a very gradual process extending over a number of years, and that only the very basic tools of language, such as frequently used prepositions, verb forms, and pronouns, tend to have been acquired by the age of five years. Other studies, which confirm that a considerable number of aspects of syntax are not well established in five-year-olds and that many interacting linguistic and general cognitive problems are involved include: in relation to English — Bever (1970), Brown (1971), Cromer (1971), Gaer (1969), Hayhurst (1967), Limber (1973), Maratsos (1976), Sheldon (1974), and Weil and Stenning (1978); in relation to French — Ferreiro (1971), Kail (1975), and Sinclair and Ferreiro (1970); in relation to Swiss German — Caprez, Sinclair, and Studer (1971); and in relation to Spanish — Ferreiro (1974).

Nouns, verbs and verb tenses, and function words including prepositions and pronouns have been regarded as among the most basic classes of words in the child's syntactic repertoire (Berko, 1958; Chapman, 1979; Clark, 1973; Slobin, 1967) and thus need to be assessed. A test of negatives has been included, although some studies suggest (Bloom, 1970) that many children have no problems with these syntactic constructions. The greater complexity of negative linguistic constructions for children from non-English language backgrounds has been noted in the literature (Bloom, 1970; Brown, 1973; Wode, 1977).

The remaining area covered in the syntax tests of the ACER Early School Series is listening comprehension. Comprehension is regarded as one of the categories of language performance (McNeill, 1966) and has been defined as 'understanding', a concept which in the area of syntax implies the understanding of the structure of sentences (Smith, 1973). The literature does not make it clear whether syntactic structure is, in fact, the most important element in the ability to understand a sentence. None the less, practice in general comprehension can be expected to be of benefit to the child, not only in relation to general communication, but also in relations to his or her ability to learn to read.

While some reading experts have assumed that word recognition and comprehension are highly related (Durrell, 1968), others have argued that the relationship is less substantial (e.g. Kendall and Hood, 1979; Wiener and Cromer, 1967). Both Cromer (1970) and Guthrie (1973) have concluded that there may be two types of disabled readers, one type primarily deficient in word recognition skills and the other primarily deficient in comprehension skills.

It has been shown that readers with good comprehension but poor recognition abilities 'make errors' that more often are contextually than graphically constrained' (Kendall and Hood, 1979, p.42). Cromer (1970) found that readers with poor comprehension but adequate word recognition skills fail to make meaningful relationships among words. The latter ability would seem to be extremely important, and was thus regarded as highly desirable to include a comprehension test in the present pre-reading battery.

The tests of prepositions, verb tense, pronouns, and negation were designed to
provide the teacher with an indication of the child's comprehension of basic syntactic structures in a simplified form so that any failure to comprehend a particular structure is easily noticeable by the teacher. The Comprehension Test requires a greater degree of attention, auditory memory, and the integration of various pieces of information. This test, among other uses previously discussed, should provide a means for the identification of those children who, while they can understand a single sentence, find the concentrated listening and interpretation of more complex materials such as stories, instructions at school, etc., much more difficult.

Description

Syntactic Structures are assessed in a total of 43 items on the basis of the following separate tests:

- Prepositions Test (9 items)
- Verb Tense Test (9 items)
- Pronouns Test (11 items)
- Negation Test (8 items)
- Comprehension Test (6 items)

Each item in all tests consists of a stimulus sentence or brief stimulus story read by the tester, to which the child is required to respond by indicating which one of an array of four drawings best represents the stimulus passage. The recommended response for group administration of the tests is that the child place a cross on the correct picture.

Where possible, the items for each test were produced in such a way that they would fit into a framework of story content which is unique to the particular test. For example, all items of the Prepositions Test relate to the activities of Sooty, the cat. The Verb Tense Test is concerned with a family of bears.

The stimulus material for the Pronouns Test relates to activities of a human family. The items for the Negation and Comprehension Tests are less uniform in content. The reason for this was that preliminary item trials had shown that a uniform story framework for these tests was more likely to provide the testee with extraneous cues, thus decreasing the validity of the assessment of the target skills of syntactic structure.

Some teachers may wish to use one or only some of the tests. This would depend on the purpose of assessment. The tests are suitable for use as achievement tests to assess the efficacy of teaching certain grammatical constructions. Wherever the aim is to obtain a more general impression of a child's or a group's language ability, it is recommended that all five tests be administered. The tests have been designed to be as short as possible so as to provide the teacher with a broad sample of the child's language functioning with minimal expenditure of time. It is possible to administer all five receptive language tests in one session. If this procedure is desired, it is recommended that the tests be administered in the order in which they are being presented in this Handbook, namely: Prepositions, Verb Tense, Pronouns, Negation, Comprehension. However, as each test forms a self-contained unit, it is not essential that the above order be followed. Wherever the tests are administered over a number of sessions, the order of administration...
depends on the preferences of the teacher. All tests are suitable for retesting purposes.

The drawings are clear and simple and contain little detail. Even individuals with minor problems in visual acuity are not expected to experience difficulties with them.

Development
The need to code the test items as pictures placed some constraints on the choice of stimulus constructions. For example, only prepositions and verb forms which could be conveyed by relatively simple drawings could be included in the tests. All items were initially tried out in one-to-one testing situations. Trial testing took place in groups of between five and ten children.

Research data on the tests were obtained from 217 school beginners after they had been at school for almost five months. The sample was drawn from five Victorian schools. Two of these schools were situated in the inner suburban area of Melbourne, which has a high incidence of children from non-English language backgrounds. Language backgrounds in this group included Italian, Greek, Macedonian, Portuguese, Serbo-Croatian, Spanish, Turkish, and Vietnamese. One school was from a high socio-economic suburb of Melbourne, one from a high socio-economic area in a provincial city, and one from a country area. Testing was carried out in two sessions on consecutive days for each child.

Administration and Scoring
The time required for administration of all five tests of syntax depends on the maturity and possible handicaps of the testee, and on the size of the group to whom the tests are being administered. During trial testing none of the tests required more than 10 to 15 minutes, and it was found that on the Comprehension Test few children required more than one minute per item. In fact, it was found that the performances of slow responders tended not to improve when more time was allowed. In the other tests, most children required no more than 30 seconds to complete each item. All tests included in the ACER Early School Series are designed to be power tests, not speed tests.

As previously noted, the language tests can be administered either in one session or over a period of time. If the latter procedure is adopted, it is suggested that each one of the syntax tests be administered as a total test in one session. Some of the other tests contained in the ACER Early School Series (e.g. the Auditory Discrimination Test and Recognition of Initial Consonant Sounds Test) were designed to be administered either as a whole or in part. Each one of the tests of syntactic structure is so short that administration in one session should not prove to be difficult even with very immature or otherwise handicapped individuals.

Materials required for each test consist of the Directions for Administration for the specific test which is to be administered, a test booklet, cardboard marker, crayon or pencil for each child, and a spare test booklet which the tester uses for demonstration purposes. The Directions for Administration contain general and detailed instructions for the specific tests. The directions are in simple English and should be read out in full for each test item.
Scoring

A score key is provided with each set of tests. Scoring has been made as simple and as fast as possible. One point is given for each correct item. Possible points for each item are zero (0) or one (1). The raw score for each test is provided by the sum of correctly completed items. Tables 8.4 and A.3 allow for the conversion of raw scores to Rasch-scaled language ability scores.

Research Results

In this section, a summary is provided of the results obtained on the tests of syntactic structures and the Comprehension Test from a sample of 217 school beginners, described in the section on Development. All five tests were administered to the children approximately five months after the beginning of the school year. These research results are presented here to provide comparative data which may help users of the tests in the interpretation of their test results.

Tables 8.1 to 8.5 describe the performance of the above sample on each of the syntax tests respectively. Differences in the number of children tested (N) resulted from the fact that the tests were administered in two sessions and some children were absent for one of these.

Comparing the children's performance relating to different aspects of syntax, a number of interesting observations were made. None of the children in the sample completed all of the items in the Prepositions Test correctly. Across the other four tests of syntactical structures, on the average less than half of the children in the total sample (i.e. \( \bar{X} = 43 \) per cent, \( SD = 14 \) per cent) completed all items correctly. This proportion rises to an average of 53 per cent (\( SD = 14 \) per cent) for children whose native language is English, but an average of only 21 per cent (\( SD = 15.0 \) per cent) of children from non-English language backgrounds tended to complete all items of the Verb Tense, Pronouns, Negation, and Comprehension Tests correctly.

The Negation Test was found to be the easiest test by all groups. Seventy-three per cent of native speakers of English, but only 44 per cent of the non-English language background children completed all items of this test correctly. The average item facilities were 94 per cent (SD = 3.9 per cent) and 80.2 per cent (SD = 8.3 per cent) for the two groups respectively.

The Pronouns Test was found to be the most difficult of the syntax tests for the total sample, as well as for the groups of native speakers of English, and non-English native language when analysed separately. One third of the total sample completed all items of this test correctly. Analyses according to language background showed 42 per cent of the native speakers of English but only 19 per cent of the non-English background group to complete all items correctly. The average item facilities for the two groups were 88 per cent (SD = 8 per cent) and 66 per cent (SD = 6 per cent) respectively. The orders of difficulty of the syntax tests from easiest to most difficult for the total sample, for native speakers of English, and for children from non-English language backgrounds are shown in Table 8.6.

While language background does not appear to have influenced the order of difficulty of the tests very much, the differences in relative performance between the tests for non-English and English language background children may contain useful diagnostic hypotheses.
### Table 8.1  Prepositions Test: Performance of 217 School Beginners Five Months after the Beginning of the School Year

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-English</td>
</tr>
<tr>
<td><strong>Total sample</strong></td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>8 (41%)</td>
</tr>
<tr>
<td>Median</td>
<td>7</td>
</tr>
<tr>
<td>Mean</td>
<td>6.5</td>
</tr>
<tr>
<td>SD</td>
<td>1.7</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>83.9%</td>
</tr>
<tr>
<td>SD</td>
<td>11.5%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.68</td>
</tr>
<tr>
<td>N</td>
<td>217</td>
</tr>
</tbody>
</table>

### Table 8.2  Verb Tense Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-English</td>
</tr>
<tr>
<td><strong>Total sample</strong></td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>9 (38%)</td>
</tr>
<tr>
<td>Median</td>
<td>8</td>
</tr>
<tr>
<td>Mean</td>
<td>7.7</td>
</tr>
<tr>
<td>SD</td>
<td>1.6</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>85.5%</td>
</tr>
<tr>
<td>SD</td>
<td>14.0%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.71</td>
</tr>
<tr>
<td>N</td>
<td>214</td>
</tr>
</tbody>
</table>
### Table 8.3 Pronoun Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>Non-English</td>
</tr>
<tr>
<td>Mode</td>
<td>11 (33%)</td>
</tr>
<tr>
<td>Median</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>8.8</td>
</tr>
<tr>
<td>SD</td>
<td>2.6</td>
</tr>
<tr>
<td>SE</td>
<td>0.2</td>
</tr>
<tr>
<td>Average item facility</td>
<td>80.1%</td>
</tr>
<tr>
<td>SD</td>
<td>8.8%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.82</td>
</tr>
<tr>
<td>N</td>
<td>214</td>
</tr>
</tbody>
</table>

### Table 8.4 Negation Test: Performance of 216 School Beginners Five Months after the Beginning of the School Year

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>Non-English</td>
</tr>
<tr>
<td>Mode</td>
<td>8 (64%)</td>
</tr>
<tr>
<td>Median</td>
<td>8</td>
</tr>
<tr>
<td>Mean</td>
<td>7.3</td>
</tr>
<tr>
<td>SD</td>
<td>1.3</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>90.7%</td>
</tr>
<tr>
<td>SD</td>
<td>5.2%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.72</td>
</tr>
<tr>
<td>N</td>
<td>216</td>
</tr>
</tbody>
</table>
Although the performance of the non-English language background group was far poorer on all tests than the performance of the groups of native speakers of English, the relative ranking of difficulty for some individual items was shown to differ for the two groups. It is noted that some items seemed relatively easier than others for the non-English background group, but not for the native speakers of English. Reasons for this are unclear. A possible explanation is that certain items may elicit different skills from native speakers than from second or subsequent language learners. Other explanations might be related to differences in effort and attention in some non-English background children, and to other compensatory mechanisms. The performance of the non-English group was certainly more varied than that of the group of native speakers, as reflected by the consistently larger standard deviations obtained for the former group in all tests.

Tables 8.7 to 8.11 show for each test, the difficulty found by children in the total sample, the non-English background group, and the native speakers of English for each item. The facility of each item is expressed by the percentage of children who completed the item correctly. The order of items for the final version of each test was determined by the performance of the total sample.

### Table 8.5 Comprehension Test: Performance of 214 School Beginners Five Months after the Beginning of the School Year

<table>
<thead>
<tr>
<th>Mode</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
<th>Average item facility</th>
<th>SD</th>
<th>KR 20</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>81.5%</td>
<td>8.1%</td>
<td>0.55</td>
<td>214</td>
</tr>
</tbody>
</table>

Obtained scores

<table>
<thead>
<tr>
<th>Total sample</th>
<th>Home-language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-English</td>
</tr>
<tr>
<td>Mode</td>
<td>6 (39%)</td>
</tr>
<tr>
<td>Median</td>
<td>5</td>
</tr>
<tr>
<td>Mean</td>
<td>4.8</td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
</tr>
<tr>
<td>SE</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Although the performance of the non-English language background group was far poorer on all tests than the performance of the groups of native speakers of English, the relative ranking of difficulty for some individual items was shown to differ for the two groups. It is noted that some items seemed relatively easier than others for the non-English background group, but not for the native speakers of English. Reasons for this are unclear. A possible explanation is that certain items may elicit different skills from native speakers than from second or subsequent language learners. Other explanations might be related to differences in effort and attention in some non-English background children, and to other compensatory mechanisms. The performance of the non-English group was certainly more varied than that of the group of native speakers, as reflected by the consistently larger standard deviations obtained for the former group in all tests.

### Table 8.6 Tests of Syntax: Order of Difficulty from Easiest to Most Difficult

<table>
<thead>
<tr>
<th>Test</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total sample</td>
</tr>
<tr>
<td>Negation</td>
<td>(64%; 91%)</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>(38%; 86%)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>(39%; 81%)</td>
</tr>
<tr>
<td>Prepositions</td>
<td>(0%; 84%)</td>
</tr>
<tr>
<td>Pronouns</td>
<td>(33%; 80%)</td>
</tr>
</tbody>
</table>

a The percentages in brackets indicate the proportion of the sample completing the total test correctly, and the average facility of the items in each test.
### Table 8.7  Prepositions Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>95</td>
<td>91</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>92</td>
<td>87</td>
<td>94</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>76</td>
<td>96</td>
</tr>
<tr>
<td>5</td>
<td>87</td>
<td>66</td>
<td>96</td>
</tr>
<tr>
<td>6</td>
<td>79</td>
<td>47</td>
<td>93</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>37</td>
<td>89</td>
</tr>
<tr>
<td>8</td>
<td>72</td>
<td>51</td>
<td>81</td>
</tr>
<tr>
<td>9</td>
<td>68</td>
<td>56</td>
<td>74</td>
</tr>
<tr>
<td>N</td>
<td>217</td>
<td>68</td>
<td>149</td>
</tr>
</tbody>
</table>

### Table 8.8  Verb Tense Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96</td>
<td>90</td>
<td>98</td>
</tr>
<tr>
<td>2</td>
<td>94</td>
<td>86</td>
<td>98</td>
</tr>
<tr>
<td>3</td>
<td>93</td>
<td>83</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>93</td>
<td>78</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>90</td>
<td>68</td>
<td>99</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>73</td>
<td>95</td>
</tr>
<tr>
<td>7</td>
<td>87</td>
<td>79</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>59</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>51</td>
<td>24</td>
<td>62</td>
</tr>
<tr>
<td>N</td>
<td>214</td>
<td>63</td>
<td>151</td>
</tr>
</tbody>
</table>

### Table 8.9  Pronouns Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>88</td>
<td>67</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>87</td>
<td>70</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>77</td>
<td>91</td>
</tr>
<tr>
<td>4</td>
<td>86</td>
<td>69</td>
<td>93</td>
</tr>
<tr>
<td>5</td>
<td>86</td>
<td>63</td>
<td>95</td>
</tr>
<tr>
<td>6</td>
<td>84</td>
<td>66</td>
<td>91</td>
</tr>
<tr>
<td>7</td>
<td>81</td>
<td>53</td>
<td>91</td>
</tr>
<tr>
<td>8</td>
<td>81</td>
<td>66</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>73</td>
<td>56</td>
<td>81</td>
</tr>
<tr>
<td>10</td>
<td>67</td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td>11</td>
<td>62</td>
<td>36</td>
<td>73</td>
</tr>
<tr>
<td>N</td>
<td>214</td>
<td>64</td>
<td>150</td>
</tr>
</tbody>
</table>
### Table 8.10 Negation Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>98</td>
<td>92</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>94</td>
<td>83</td>
<td>99</td>
</tr>
<tr>
<td>3</td>
<td>93</td>
<td>86</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>93</td>
<td>86</td>
<td>95</td>
</tr>
<tr>
<td>5</td>
<td>91</td>
<td>77</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>90</td>
<td>80</td>
<td>95</td>
</tr>
<tr>
<td>7</td>
<td>86</td>
<td>68</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>81</td>
<td>70</td>
<td>87</td>
</tr>
<tr>
<td>N</td>
<td>216</td>
<td>66</td>
<td>150</td>
</tr>
</tbody>
</table>

### Table 8.11 Comprehension Test: Item Difficulties Expressed as Percentage Correct

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>86</td>
<td>74</td>
<td>91</td>
</tr>
<tr>
<td>3</td>
<td>79</td>
<td>72</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>79</td>
<td>62</td>
<td>86</td>
</tr>
<tr>
<td>5</td>
<td>79</td>
<td>63</td>
<td>85</td>
</tr>
<tr>
<td>6</td>
<td>71</td>
<td>52</td>
<td>79</td>
</tr>
<tr>
<td>N</td>
<td>214</td>
<td>65</td>
<td>149</td>
</tr>
</tbody>
</table>
### Table 8.12 Prepositions Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SE_{Di})

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SE_{Di}</td>
<td>Di</td>
</tr>
<tr>
<td>1</td>
<td>56.1</td>
<td>1.0</td>
<td>57.7</td>
</tr>
<tr>
<td>2</td>
<td>54.8</td>
<td>1.0</td>
<td>54.9</td>
</tr>
<tr>
<td>3</td>
<td>54.6</td>
<td>1.0</td>
<td>53.9</td>
</tr>
<tr>
<td>4</td>
<td>52.1</td>
<td>1.0</td>
<td>52.8</td>
</tr>
<tr>
<td>5</td>
<td>48.6</td>
<td>1.1</td>
<td>50.1</td>
</tr>
<tr>
<td>6</td>
<td>46.7</td>
<td>1.3</td>
<td>46.9</td>
</tr>
<tr>
<td>7</td>
<td>45.4</td>
<td>1.4</td>
<td>42.3</td>
</tr>
<tr>
<td>8</td>
<td>41.8</td>
<td>1.8</td>
<td>41.3</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td>6.0</td>
<td>50.0</td>
</tr>
<tr>
<td>SD</td>
<td>5.2</td>
<td>5.5</td>
<td>5.2</td>
</tr>
<tr>
<td>N</td>
<td>128</td>
<td>128</td>
<td>62</td>
</tr>
</tbody>
</table>

*a Item 1 was excluded because Rasch analyses reject items which all subjects in the sample complete correctly or incorrectly.

### Interpretation of Scores

Tables 8.12 to 8.16 show the estimates of Rasch item difficulties (Di) and standard errors (SE_{Di}) for the difficulty estimates for each item in the total sample, the group of non-English background children, and the sample of native speakers of English.
## Table 8.14 Pronouns Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SE_{Di})

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SE_{Di}</td>
<td>Di</td>
</tr>
<tr>
<td>1</td>
<td>57.0</td>
<td>1.0</td>
<td>57.1</td>
</tr>
<tr>
<td>2</td>
<td>55.0</td>
<td>1.0</td>
<td>53.1</td>
</tr>
<tr>
<td>3</td>
<td>52.9</td>
<td>1.0</td>
<td>51.8</td>
</tr>
<tr>
<td>4</td>
<td>50.0</td>
<td>1.1</td>
<td>51.4</td>
</tr>
<tr>
<td>5</td>
<td>50.0</td>
<td>1.1</td>
<td>49.3</td>
</tr>
<tr>
<td>6</td>
<td>48.8</td>
<td>1.1</td>
<td>48.9</td>
</tr>
<tr>
<td>7</td>
<td>47.9</td>
<td>1.2</td>
<td>48.9</td>
</tr>
<tr>
<td>8</td>
<td>47.7</td>
<td>1.2</td>
<td>48.4</td>
</tr>
<tr>
<td>9</td>
<td>47.4</td>
<td>1.2</td>
<td>48.0</td>
</tr>
<tr>
<td>10</td>
<td>46.9</td>
<td>1.3</td>
<td>47.5</td>
</tr>
<tr>
<td>11</td>
<td>46.4</td>
<td>1.3</td>
<td>45.5</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>SD</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>144</td>
<td>57</td>
<td>87</td>
</tr>
</tbody>
</table>

## Table 8.15 Negation Test: Rasch-scaled Item Difficulty Estimates (Di) and Standard Errors (SE_{Di})

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SE_{Di}</td>
<td>Di</td>
</tr>
<tr>
<td>1</td>
<td>56.8</td>
<td>1.2</td>
<td>55.5</td>
</tr>
<tr>
<td>2</td>
<td>54.3</td>
<td>1.2</td>
<td>54.9</td>
</tr>
<tr>
<td>3</td>
<td>51.2</td>
<td>1.4</td>
<td>51.9</td>
</tr>
<tr>
<td>4</td>
<td>50.8</td>
<td>1.4</td>
<td>50.5</td>
</tr>
<tr>
<td>5</td>
<td>49.2</td>
<td>1.5</td>
<td>49.1</td>
</tr>
<tr>
<td>6</td>
<td>48.8</td>
<td>1.5</td>
<td>47.5</td>
</tr>
<tr>
<td>7</td>
<td>47.3</td>
<td>1.7</td>
<td>47.5</td>
</tr>
<tr>
<td>8</td>
<td>41.4</td>
<td>2.7</td>
<td>43.0</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>SD</td>
<td>4.6</td>
<td></td>
<td>4.1</td>
</tr>
<tr>
<td>N</td>
<td>77</td>
<td>36</td>
<td>41</td>
</tr>
</tbody>
</table>
None of the tests showed statistically significant differences in the item difficulties between the non-English language background group and the native speakers of English.

Tables 8.17 to 8.21 contain the Rasch ability estimates (Ai) and the standard errors (SEA) for each estimate for the tests.

Inspection of Tables 8.17 to 8.21 shows that the ability estimates obtained from the raw scores show very close agreement between the non-English and the English language background groups and the total sample. This provides evidence for the stability of the ability estimates obtained by the Rasch method. In this type of measurement model, the ability estimate obtained from the raw score on a test is quite independent of the characteristics of the sample of children from which it was obtained. The ability estimate for an individual will thus be the same, no matter in which group of tests he or she was assessed.
## TESTS OF SYNTAX

### Table 8.18 Verb Tense Test: Rasch Ability Estimates (Ai) and Standard Errors (SE<sub>AI</sub>)

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
</tr>
<tr>
<td>1</td>
<td>37.5</td>
<td>6.4</td>
<td>37.1</td>
<td>6.5</td>
<td>36.0</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>42.5</td>
<td>4.8</td>
<td>42.2</td>
<td>5.0</td>
<td>41.6</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>45.8</td>
<td>4.3</td>
<td>45.7</td>
<td>4.4</td>
<td>45.3</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>48.7</td>
<td>4.0</td>
<td>48.6</td>
<td>4.1</td>
<td>51.5</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>51.3</td>
<td>4.0</td>
<td>51.4</td>
<td>4.1</td>
<td>51.5</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>54.3</td>
<td>4.3</td>
<td>54.3</td>
<td>4.4</td>
<td>54.7</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>57.7</td>
<td>4.8</td>
<td>57.8</td>
<td>5.0</td>
<td>58.4</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>62.5</td>
<td>6.4</td>
<td>62.9</td>
<td>6.5</td>
<td>64.0</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>133</td>
<td>133</td>
<td>55</td>
<td>55</td>
<td>78</td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8.19 Pronouns Test: Rasch Ability Estimates (Ai) and Standard Errors (SE<sub>AI</sub>)

<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
</tr>
<tr>
<td>1</td>
<td>38.3</td>
<td>5.2</td>
<td>38.7</td>
<td>5.2</td>
<td>37.6</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>42.5</td>
<td>3.9</td>
<td>42.6</td>
<td>3.9</td>
<td>41.9</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>45.1</td>
<td>3.4</td>
<td>45.2</td>
<td>3.4</td>
<td>44.7</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>47.2</td>
<td>3.1</td>
<td>47.2</td>
<td>3.1</td>
<td>47.0</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>49.1</td>
<td>3.0</td>
<td>49.1</td>
<td>3.0</td>
<td>49.0</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>50.9</td>
<td>3.0</td>
<td>50.9</td>
<td>3.0</td>
<td>51.0</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>52.8</td>
<td>3.1</td>
<td>52.8</td>
<td>3.1</td>
<td>53.0</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>54.9</td>
<td>3.4</td>
<td>54.8</td>
<td>3.4</td>
<td>55.3</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>57.5</td>
<td>3.9</td>
<td>57.4</td>
<td>3.9</td>
<td>58.1</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>61.5</td>
<td>5.2</td>
<td>61.3</td>
<td>5.2</td>
<td>62.4</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>144</td>
<td>144</td>
<td>57</td>
<td>57</td>
<td>87</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8.20 Negation Test: Rasch Ability Estimates (Ai) and Standard Errors (SE<sub>AI</sub>)

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Total sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
<td>SE&lt;sub&gt;AI&lt;/sub&gt;</td>
<td>Ai</td>
</tr>
<tr>
<td>1</td>
<td>39.7</td>
<td>5.7</td>
<td>40.0</td>
<td>5.7</td>
<td>38.8</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>44.2</td>
<td>4.4</td>
<td>44.3</td>
<td>4.4</td>
<td>43.7</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>47.3</td>
<td>3.9</td>
<td>47.4</td>
<td>3.9</td>
<td>47.1</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>50.0</td>
<td>3.8</td>
<td>50.0</td>
<td>3.8</td>
<td>50.0</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>52.7</td>
<td>3.9</td>
<td>52.6</td>
<td>3.9</td>
<td>52.9</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>55.8</td>
<td>4.4</td>
<td>55.7</td>
<td>4.4</td>
<td>56.3</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>60.3</td>
<td>5.7</td>
<td>60.0</td>
<td>5.7</td>
<td>61.2</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>77</td>
<td>77</td>
<td>36</td>
<td>36</td>
<td>41</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>
THE TESTS

Table 8.21  Comprehension Test: Rasch Ability Estimates (AI) and Standard Errors (SEAI)

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SEAI</td>
<td>Ai</td>
</tr>
<tr>
<td>1</td>
<td>41.6</td>
<td>5.7</td>
<td>41.3</td>
</tr>
<tr>
<td>2</td>
<td>46.6</td>
<td>4.5</td>
<td>46.2</td>
</tr>
<tr>
<td>3</td>
<td>50.0</td>
<td>4.2</td>
<td>50.0</td>
</tr>
<tr>
<td>4</td>
<td>53.6</td>
<td>4.5</td>
<td>53.8</td>
</tr>
<tr>
<td>5</td>
<td>58.4</td>
<td>5.7</td>
<td>58.7</td>
</tr>
<tr>
<td>N</td>
<td>130</td>
<td>130</td>
<td>53</td>
</tr>
</tbody>
</table>

Pearson's 'r' correlation coefficients between the item difficulties and the ability estimates based on raw scores for the two samples (i.e. non-English and English language background) were computed and are shown in Table 8.22.

A perfect correlation between the ability estimates made on the basis of the raw scores for non-English and English language groups shows that the conversion from a specific raw score to a Rasch ability estimate is constant across samples and testing situations.

Implications

The more complex a behaviour the more difficult it is to assess its validity and reliability:

The measurements we make can never capture all the complexity, nuances, and diverse manifestations of (language) behavior. We can only selectively sample some subset of that behavior in certain situations and, from the sample, infer availability of the full range of the behaviour in all the situations which call for it. (Lamberts, 1979, p.272)

Each of the tests of syntax contained in the ACER Early School Series assesses the child's understanding of a particular set of linguistic constructs by presenting a small sample of examples from that domain. For example, the set of nine frequently occurring prepositions or that of 11 pronouns are samples only of the

Table 8.22  Tests of Syntactic Structure: Correlations between Item Difficulties (Di) and between Ability Estimates Obtained in Two Independent Samples

<table>
<thead>
<tr>
<th>Test</th>
<th>Correlation between non-English and English language background groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
</tr>
<tr>
<td>Prepositions</td>
<td>.90</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>.97</td>
</tr>
<tr>
<td>Pronouns</td>
<td>.95</td>
</tr>
<tr>
<td>Negation</td>
<td>.89</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.98</td>
</tr>
</tbody>
</table>
child's repertoire in these domains. The individual's performance on each of the tests leads the teacher to infer that the child has reached an understanding of the basic syntactic structures involved.

Another important consideration needs to be raised at this point. Language competence is comprised of both comprehension and production, that is, reception and expression. In other words the individual's use of language involves the ability to understand messages produced by others as well as the ability to emit vocal messages.

In accordance with the widely held notion that children tend to understand much more than they themselves can express, linguistics experts have argued that receptive language tests best reveal the individual's language ability (Bellugi-Klima, 1968). These researchers have suggested that comprehension tasks place a considerably greater burden on knowledge of language, and thus tend to provide better measures of competence, than verbal, expression. Clark, Hutcheson, and Van Buren (1974) provided strong arguments for the view that to talk may be easier for the young child than understanding the talk of others (i.e. comprehension). The reason given for this is that, if the child is talking, 'he will be remarking on those aspects of the situation which for him are prominent' (Clark et al., 1974, p.49). However, when required to comprehend the words spoken by another person, the child 'must redirect his attention to attain their view of the situation' (Clark et al., 1974, p.49). It is suggested that this adjustment of attention required in the latter instance may often be difficult for children to make (Clark et al., 1974, Lamberts, 1979).

Those interested in the psycholinguistic debate of whether and, if so, to what degree children's ability to comprehend syntactic and other linguistic structures is more or less advanced than their ability to express themselves are referred to the work by Ingram (1974, 1976). However, no matter which side of this theoretical issue one might support, it must be accepted that an assessment of the child's competence in language should involve samples of both receptive and expressive language.

A previously mentioned major constraint placed on the construction of the ACER Early School Series was that all tests had to provide the possibility of being administered as paper and pencil tests. It is important, therefore, for the teacher to complement the information obtained on the basis of the receptive language tests contained in the ACER Early School Series by samples of the child's ability in expressive language. In this context a warning expressed by Lamberts (1979) may be useful:

Teachers must become more careful not to overestimate the child's linguistic ability on the basis of correct responses to instructions or other verbalizations which arise in the context of daily activities. Children learn to respond to the whole context of messages addressed to them in more or less routine contexts. On the basis of past experiences, they form expectancies as to the most likely verbal message given to them in such contexts and thus may be able to respond appropriately even if they do not fully understand the message itself. Adults typically accompany their instructions with many non-verbal cues in the form of bodily or facial gestures which will aid the child's interpretation. (p.273)
To assess an individual’s ability to respond to language rather than to the total set of cues contained in the communication situation, the tester needs to delete all non-language cues which the child might utilize to interpret the verbal message (e.g. Chapman, 1972; Clark et al., 1974; Lamberts, 1979; Mittler, 1976).

Berko (1958) developed an ingenious technique of testing which has been used both with retarded and normal children, and which can easily be adapted by the teacher. The method consists of presenting the child with pairs of drawings representing imaginary figures which have been given a name (e.g. ‘wug’). The test is made up of 27 items covering common syntactic structures such as plurals, pronouns, comparative/superlative forms of adjectives, past tense, etc. The child is shown the pictures and required to give their name, in other words to say the required grammatical form.

A number of American standardized tests of expressive language exist. Like Berko’s (1958) method, all these instruments require individual testing in one-to-one situations. Examples of such tests are Elicited Language Inventory (Carrow, 1974); Michigan Picture Inventory (Lerea, 1962); Northwestern Syntax Screening Test (Lee, 1969); Sentence Imitation Test (Menyuk, 1963; 1964); Sentence Repetition Test (Anastasiow and Hanes, 1974); Stephens Oral Language Screening Test—SOLST (Stephens, 1977).

Intervention strategies and remediation procedures will include practice in simple and increasingly more difficult forms of syntax and vocabulary.

To provide a ready starting point for intervention, each test item in the syntax tests of the ACER Early School Series has been designed in such a way that it forms a starting point for the development of a teaching sequence on the particular aspect of language under scrutiny.

In the array of pictures making up each item, the picture representing the correct response is accompanied by three distractors which illustrate related structures or concepts. The child’s test performance provides diagnostic information which indicates where special help is required. The distractor pictures in each item suggest to the teacher initial extensions and alternatives. For example, the item ‘on’ is accompanied by pictures which illustrate the concepts ‘under’, ‘beside’ and ‘away from’; the picture which illustrates ‘over’ is accompanied by pictures which illustrate ‘on’, ‘in front of’, ‘beside’. Thus, in addition to using the tests to pinpoint possible language difficulties, the teacher can find a wealth of related material which can be incorporated into the language program.

Because the tests require only a marking or pointing response, they allow the teacher to make some assessment of the language comprehension of those children who are extremely timid, who have a severe speech defect, who are aphasic, or whose spoken English is not yet sufficiently fluent for them to communicate freely.
WORD KNOWLEDGE TEST

Rationale
A careful review of the literature describing educational research into word knowledge showed two major focuses, namely studies aimed at the estimation of total vocabulary size (e.g. Barbe and Abbott, 1975; Dolch, 1936; Lorge and Chall, 1963; Rinsland, 1945; Smith, 1941) and research investigating correlations between word knowledge and performance on intelligence (Dunn, 1965; Jordan, 1978; Zigler, Abelson, and Seitz, 1973) and school achievement tests (e.g. Loban, 1963).

It has been pointed out that, because of the need to keep the length of tests to a reasonable limit, and as a result of the procedure of treating the meaning of a word as a discrete ‘bit’ of information which the child possesses or does not possess, word knowledge tests tend to provide a rather limited if not invalid assessment of the child’s ability to use and comprehend words. Cazden (1971), for example, reminds the users of such tests as the Peabody Picture Vocabulary Test that vocabulary growth is just as much the growth and extended enrichment of meanings for words already in the child’s repertoire as the addition of new words.

The aim of the *Word Knowledge Test* contained in the ACER Early School Series is neither to estimate the size of an individual’s vocabulary, nor to assess a wide range of comprehension ability. Instead the *Word Knowledge Test* was designed to identify those children who may be unable to comprehend a sample of English words which many teachers in Australia would expect to be part of the vocabulary of a five-year-old child.

Individuals for whom the test is intended include the following groups:

(i) school beginners whose understanding of the English language is extremely limited, and for whom it may be helpful to identify some words which are already known, as a starting point for communication and further vocabulary development. Many of these children would be recent arrivals from non-English-speaking countries, and children whose home language is predominantly other than English;

(ii) children from non-English language backgrounds who, in spite of
THE TESTS

Apparently adequate communication skills, require a special language program because of serious gaps in their English vocabulary.

(iii) children who are suspected of having severe language limitations, either through maturational or social causes, and who will require special help in acquiring those language skills which are basic prerequisites to school learning (particularly to learning to read) and intellectual and social development more generally.

(iv) children with severe speech problems, and intellectually handicapped children and adults.

It is suggested that the *Word Knowledge Test* may also be useful as an introductory instrument for speech pathologists. Aphasic or extremely withdrawn individuals and children who are too shy to communicate verbally with either the teacher or peers can be tested with this test which does not require any verbalization on the part of the testee. In special circumstances (e.g. in cases of physical disability) the response may be indicated by pointing, or elicited directly by the tester who then records the response for the testee. The *Word Knowledge Test* may also be used by kindergarten teachers and by parents.

**Description**

The *Word Knowledge Test* is a 38-item basal vocabulary test. The child is required to respond to a stimulus word spoken by the tester by selecting the correct picture from an array of four conceptually related black and white line drawings. The child records his or her response to the stimulus word by placing a cross (X) on the correct picture in the expendable test booklet.

The drawings are simple and easily discernible. The four pictures pertaining to one item are arranged in one row.

**Development**

The basic item pool for the development of the *Word Knowledge Test* was provided by vocabulary lists and sub-tests from a large range of currently used early childhood intelligence and achievement tests and *A Word List for Australian Schools* (Radford, 1960). Nine hundred and sixty words from these sources were listed alphabetically and submitted to a representative sample of 120 teachers of school beginners in government and Catholic schools from three Australian States. The teachers were asked to select those words which they would expect all or most five-year-old children to be able to recognize and identify pictorially. Words which were judged to be in this category by at least 80 per cent of teachers were re-listed, and from these words the stimulus words were derived by random selection. Where possible, words from the list were used as distractors.

To increase the validity of this type of assessment of the child’s understanding of each stimulus word, the three distractor illustrations are from the same category of objects. For example, the target word “bike” is grouped with other nouns from the category “transport” (i.e. a train, a truck, and a car). The target word “hammer” is grouped with distractors from the category of “tool”. Other categories include animals, parts of the body, objects in the child’s immediate environment, and a variety of common activities.
Trial data were obtained from a sample of 215 school beginners to whom the test was administered five months after entry into school. The sample was drawn from five Victorian schools. Two schools were situated in the inner suburbs of Melbourne and had a high incidence of children from non-English-speaking backgrounds. One of the schools was situated in a high socio-economic suburb of Melbourne, one in a high socio-economic area in a provincial city, and one in a remote country area. The sample consisted of 150 children whose native language was English, and 65 children who came from families which did not use English as a language in the home.

In the final version of the *Word Knowledge Test*, the items were ordered according to their level of difficulty for the non-English background children. Table 9.2 shows the item difficulties for the English and the non-English language background groups and for the total sample. Rasch item difficulties and a means of converting raw scores into Rasch abilities are provided in Tables 9.3 and 9.4 respectively. However, Table A.3 in Appendix IV might be used more conveniently for the conversion to Rasch ability estimates.

The validity of any set of word knowledge items depends to a large degree on the circumstances and the purposes for which language is being used at any given time. The time required for administration has to be weighed against the gain in information resulting from lengthening a test. The words which made up the source list for this test have been included as Appendix I. Some teachers may wish to utilize these words, which appear frequently in tests and readers, in building up language skills. For easy reference, these words have been presented in alphabetical order.

**Administration and Scoring**

The time required for administration of all 38 items depends on the maturity and handicap of the testee and the size of the group of individuals to whom the test is being administered.

During trial testing, the total time required to administer the test to groups of five or six 5 to 5½-year-old school beginners tended to be 15 to 20 minutes. The ACER Early School Series tests are power tests, not speed tests. As pointed out in the Directions for Administration the test can be administered in sections over a number of days. Instructions for this procedure are provided immediately preceding the detailed directions for administering Example A.

Materials required consist of the Directions for Administration which contain general and detailed instructions for test administration, a test booklet, cardboard marker, and a crayon or pencil for each child.

Full instructions for the administration of the test are contained in the Directions for Administration. Wording of the instructions is simple. The tester says, for example:

- Put your marker under the first row of pictures.
- Put a cross on ARM.
- Put your marker under the next row.
- Put a cross on BALL.
Table 9.1  Word Knowledge Test: Performance of 215 School Beginners after Five Months of Schooling

<table>
<thead>
<tr>
<th>Obtained scores</th>
<th>Home language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total sample</td>
</tr>
<tr>
<td>Mode</td>
<td>38 (27%)</td>
</tr>
<tr>
<td>Median</td>
<td>37</td>
</tr>
<tr>
<td>Mean</td>
<td>35.2</td>
</tr>
<tr>
<td>SD</td>
<td>3.9</td>
</tr>
<tr>
<td>SE</td>
<td>1.1</td>
</tr>
<tr>
<td>Average item facility</td>
<td>93.1%</td>
</tr>
<tr>
<td>SD</td>
<td>6.8%</td>
</tr>
<tr>
<td>KR 20</td>
<td>0.87</td>
</tr>
<tr>
<td>N</td>
<td>215</td>
</tr>
</tbody>
</table>

Scoring

A score key is provided. One point is given for each correct item. Possible points for each item are zero (0) or one (1). The raw score on the test equals the number of correct responses. The maximum obtainable raw score on the test is 38. Tables 9.4 and A.3 allow the conversion of raw scores to Rasch developmental ability scores.

Research Results

The information contained in Tables 9.1 and 9.2 was derived from data analyses based on classical test development procedures. Table 9.1 shows the results obtained by the sample of 215 school beginners described in Chapter 3.

Table 9.1 shows that more than one third (38 per cent) of the children of English language background obtained full marks on the test. The most frequently obtained score, the average scores, and the average item facility in the non-English language group were significantly lower than the equivalent scores obtained by children of English language background. The larger standard deviations obtained in the non-English language group reflect the greater variability in the performances of this group. This finding is not unexpected if one considers that the group of non-English background children consists not only of children of quite heterogeneous language backgrounds, but that large variations in different children’s exposure to the English language must be expected quite in addition to possible maturational differences, and differences in linguistic talent.

As noted previously, the Word Knowledge Test does not purport to assess a wide range of language comprehension or linguistic ability, but rather to provide the teacher with a means of identifying which children might be linguistically disadvantaged as a result of their inability to understand frequently occurring words which most teachers might expect to be understood by five-year-olds. Conventional measures of internal test consistency are thus not really applicable, as the sets of items are not expected to provide a meaningful measure of the
Table 9.2  Word Knowledge Test: Item Difficulties Obtained from a Sample of 215 School Beginners

<table>
<thead>
<tr>
<th>Item</th>
<th>Language spoken at home</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English</td>
<td>Non-English</td>
<td>Total</td>
</tr>
<tr>
<td>1 drink</td>
<td>98.7</td>
<td>100.0</td>
<td>99.1</td>
</tr>
<tr>
<td>2 fish</td>
<td>100.0</td>
<td>98.5</td>
<td>99.5</td>
</tr>
<tr>
<td>3 horse</td>
<td>100.0</td>
<td>98.5</td>
<td>99.5</td>
</tr>
<tr>
<td>4 sheep</td>
<td>99.3</td>
<td>98.4</td>
<td>98.1</td>
</tr>
<tr>
<td>5 sleep</td>
<td>98.6</td>
<td>96.9</td>
<td>98.1</td>
</tr>
<tr>
<td>6 scissors</td>
<td>97.8</td>
<td>96.9</td>
<td>97.7</td>
</tr>
<tr>
<td>7 clock</td>
<td>99.3</td>
<td>96.9</td>
<td>98.6</td>
</tr>
<tr>
<td>8 fat</td>
<td>99.3</td>
<td>95.4</td>
<td>98.1</td>
</tr>
<tr>
<td>9 raper</td>
<td>97.8</td>
<td>95.4</td>
<td>97.2</td>
</tr>
<tr>
<td>10 finger</td>
<td>98.6</td>
<td>93.8</td>
<td>97.2</td>
</tr>
<tr>
<td>11 coat</td>
<td>93.6</td>
<td>93.8</td>
<td>93.2</td>
</tr>
<tr>
<td>12 bike</td>
<td>100.0</td>
<td>93.8</td>
<td>98.1</td>
</tr>
<tr>
<td>13 window</td>
<td>98.6</td>
<td>92.3</td>
<td>96.7</td>
</tr>
<tr>
<td>14 brother</td>
<td>97.8</td>
<td>92.3</td>
<td>96.3</td>
</tr>
<tr>
<td>15 hammer</td>
<td>100.0</td>
<td>90.8</td>
<td>97.2</td>
</tr>
<tr>
<td>16 man</td>
<td>100.0</td>
<td>90.8</td>
<td>97.2</td>
</tr>
<tr>
<td>17 pig</td>
<td>100.0</td>
<td>90.8</td>
<td>97.2</td>
</tr>
<tr>
<td>18 story</td>
<td>100.0</td>
<td>90.8</td>
<td>97.2</td>
</tr>
<tr>
<td>19 fire</td>
<td>99.3</td>
<td>89.2</td>
<td>96.3</td>
</tr>
<tr>
<td>20 grass</td>
<td>96.4</td>
<td>87.7</td>
<td>93.9</td>
</tr>
<tr>
<td>21 birthday</td>
<td>94.3</td>
<td>87.7</td>
<td>92.6</td>
</tr>
<tr>
<td>22 paint</td>
<td>100.0</td>
<td>84.6</td>
<td>94.4</td>
</tr>
<tr>
<td>23 fly</td>
<td>98.6</td>
<td>84.6</td>
<td>90.7</td>
</tr>
<tr>
<td>24 doctor</td>
<td>90.0</td>
<td>84.6</td>
<td>87.0</td>
</tr>
<tr>
<td>25 slide</td>
<td>92.1</td>
<td>83.1</td>
<td>89.8</td>
</tr>
<tr>
<td>26 push</td>
<td>99.3</td>
<td>83.1</td>
<td>94.4</td>
</tr>
<tr>
<td>27 bucket</td>
<td>99.3</td>
<td>81.5</td>
<td>94.9</td>
</tr>
<tr>
<td>28 laugh</td>
<td>95.7</td>
<td>81.5</td>
<td>91.6</td>
</tr>
<tr>
<td>29 beach</td>
<td>97.1</td>
<td>80.0</td>
<td>92.1</td>
</tr>
<tr>
<td>30 button</td>
<td>100.0</td>
<td>75.4</td>
<td>92.6</td>
</tr>
<tr>
<td>31 knife</td>
<td>98.6</td>
<td>75.4</td>
<td>91.6</td>
</tr>
<tr>
<td>32 old</td>
<td>96.4</td>
<td>75.4</td>
<td>90.7</td>
</tr>
<tr>
<td>33 happy</td>
<td>95.0</td>
<td>73.8</td>
<td>88.8</td>
</tr>
<tr>
<td>34 brush</td>
<td>95.7</td>
<td>70.8</td>
<td>88.4</td>
</tr>
<tr>
<td>35 garden</td>
<td>97.8</td>
<td>69.2</td>
<td>89.3</td>
</tr>
<tr>
<td>36 fell</td>
<td>82.9</td>
<td>55.4</td>
<td>75.3</td>
</tr>
<tr>
<td>37 nest</td>
<td>92.1</td>
<td>52.3</td>
<td>80.5</td>
</tr>
<tr>
<td>38 glass</td>
<td>79.3</td>
<td>35.4</td>
<td>67.0</td>
</tr>
</tbody>
</table>

a Expressed as item facility, i.e. percentage correct
linguistic competence of the individual. The inability to complete an item pinpoints a gap in the word knowledge of the individual, which the teacher is able to remediate. The KR 20 values have been reported in Table 9.1 because, in the case of the non-English background group and for the total sample, they reveal a moderately high degree of internal consistency in the test. The low KR 20 for the English-language group shows that the items have a low internal consistency in the assessment of the word knowledge for these children. An explanation may lie in the fact that for most of the children in this group the test is very easy; the high average scores and average item facility and the small standard deviation provide further evidence for this. It is suggested that school beginners who are native speakers of English in Australia would, in fact, understand the words contained in the test, and that any variations in the performances in this group are likely to be the result of variables other than word knowledge (e.g. lack of attention, lack of motivation, etc.).

Table 9.2 contains the difficulties for each item, separately for the English and non-English language groups and for the total sample.

**Interpretation of Scores**

For the other tests contained in the ACER Early School Series, Rasch-type item analyses were conducted as part of the test development. It was found that the rank order of Rasch item difficulties agreed basically with the rank order obtained in conventional analyses. In the case of the *Word Knowledge Test*, this was not so. As was discussed previously, standardization of the *Word Knowledge Test* is not appropriate as the test is not aimed to measure general comprehension, linguistic ability, or any other latent trait. The items in this test were thus ordered according to their difficulty level for the non-English background school beginners.

The Rasch calibration of the items of the *Word Knowledge Test* was conducted solely to link this test into the series with the other tests which make up the ACER Early School Series and to utilize, as in the case of the other tests, a common set of scale values to represent the parameters of item difficulty and test performance of individuals simultaneously. Table 9.3 shows the estimates of Rasch item difficulty ($D_i$) and standard errors ($SE_{D_i}$) of difficulties for the total sample, and separately for the non-English and English language background groups.

The correlation between the Rasch item difficulties for the non-English and English background groups was $r = 0.62$. No difference was found in the mean item difficulty between non-English and English language background children. The performances of the English language group showed greater variation between individuals than did the performances of the non-English language group. This supports the previously suggested hypothesis that test scores in this test may be influenced by different variables in the English and non-English background groups. Further research into this matter may provide interesting and useful information concerning the variables which may have influenced the performance of native speakers of English. However, the present lack of explanation for the performance of the native speakers of English, who would generally have been expected to understand the stimulus words, does not affect the usefulness of the
<table>
<thead>
<tr>
<th>Item</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Di</td>
<td>SE&lt;sub&gt;Di&lt;/sub&gt;</td>
<td>Di</td>
</tr>
<tr>
<td>1</td>
<td>64.5</td>
<td>6.9</td>
<td>66.9</td>
</tr>
<tr>
<td>2</td>
<td>64.5</td>
<td>6.9</td>
<td>62.6</td>
</tr>
<tr>
<td>3</td>
<td>64.5</td>
<td>6.9</td>
<td>61.8</td>
</tr>
<tr>
<td>4</td>
<td>54.9</td>
<td>3.5</td>
<td>58.1</td>
</tr>
<tr>
<td>5</td>
<td>59.7</td>
<td>4.9</td>
<td>57.6</td>
</tr>
<tr>
<td>6</td>
<td>59.7</td>
<td>4.9</td>
<td>56.1</td>
</tr>
<tr>
<td>7</td>
<td>56.9</td>
<td>4.0</td>
<td>56.1</td>
</tr>
<tr>
<td>8</td>
<td>54.9</td>
<td>3.5</td>
<td>56.1</td>
</tr>
<tr>
<td>9</td>
<td>53.3</td>
<td>3.1</td>
<td>56.1</td>
</tr>
<tr>
<td>10</td>
<td>54.9</td>
<td>3.5</td>
<td>54.5</td>
</tr>
<tr>
<td>11</td>
<td>47.0</td>
<td>2.0</td>
<td>53.9</td>
</tr>
<tr>
<td>12</td>
<td>54.9</td>
<td>3.5</td>
<td>53.9</td>
</tr>
<tr>
<td>13</td>
<td>52.0</td>
<td>2.9</td>
<td>53.2</td>
</tr>
<tr>
<td>14</td>
<td>50.0</td>
<td>2.5</td>
<td>53.2</td>
</tr>
<tr>
<td>15</td>
<td>52.0</td>
<td>2.9</td>
<td>32.5</td>
</tr>
<tr>
<td>16</td>
<td>52.0</td>
<td>2.9</td>
<td>52.5</td>
</tr>
<tr>
<td>17</td>
<td>53.3</td>
<td>3.1</td>
<td>52.5</td>
</tr>
<tr>
<td>18</td>
<td>50.0</td>
<td>2.5</td>
<td>50.9</td>
</tr>
<tr>
<td>19</td>
<td>50.0</td>
<td>2.5</td>
<td>50.9</td>
</tr>
<tr>
<td>20</td>
<td>50.0</td>
<td>2.5</td>
<td>49.9</td>
</tr>
<tr>
<td>21</td>
<td>46.5</td>
<td>2.0</td>
<td>48.9</td>
</tr>
<tr>
<td>22</td>
<td>48.4</td>
<td>2.2</td>
<td>48.9</td>
</tr>
<tr>
<td>23</td>
<td>45.4</td>
<td>1.9</td>
<td>48.9</td>
</tr>
<tr>
<td>24</td>
<td>45.9</td>
<td>1.9</td>
<td>48.9</td>
</tr>
<tr>
<td>25</td>
<td>48.4</td>
<td>2.2</td>
<td>47.6</td>
</tr>
<tr>
<td>26</td>
<td>47.7</td>
<td>2.1</td>
<td>47.6</td>
</tr>
<tr>
<td>27</td>
<td>47.7</td>
<td>2.1</td>
<td>46.1</td>
</tr>
<tr>
<td>28</td>
<td>45.4</td>
<td>1.9</td>
<td>46.1</td>
</tr>
<tr>
<td>29</td>
<td>45.4</td>
<td>1.9</td>
<td>46.1</td>
</tr>
<tr>
<td>30</td>
<td>44.4</td>
<td>1.7</td>
<td>44.2</td>
</tr>
<tr>
<td>31</td>
<td>45.4</td>
<td>1.9</td>
<td>44.2</td>
</tr>
<tr>
<td>32</td>
<td>44.0</td>
<td>1.7</td>
<td>41.6</td>
</tr>
<tr>
<td>33</td>
<td>42.1</td>
<td>1.5</td>
<td>41.6</td>
</tr>
<tr>
<td>34</td>
<td>44.4</td>
<td>1.7</td>
<td>41.6</td>
</tr>
<tr>
<td>35</td>
<td>42.9</td>
<td>1.6</td>
<td>37.2</td>
</tr>
<tr>
<td>36</td>
<td>39.0</td>
<td>1.3</td>
<td>37.2</td>
</tr>
<tr>
<td>37</td>
<td>39.5</td>
<td>1.3</td>
<td>37.2</td>
</tr>
<tr>
<td>38</td>
<td>38.5</td>
<td>1.3</td>
<td>37.2</td>
</tr>
<tr>
<td>X</td>
<td>50.0</td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>SD</td>
<td>6.80</td>
<td></td>
<td>7.16</td>
</tr>
</tbody>
</table>
**THE TESTS**

_The Tests_ for those children for whom it is actually intended; for example, children from non-English-speaking backgrounds whose English is severely limited or who have gaps in their English vocabulary, and individuals from an English-speaking background whose language development has been negatively affected by general developmental, physical, or social causes. The aims and usage of the _Word Knowledge Test_ have been discussed in more detail earlier in this chapter.

Table 9.4 describes the performances of the non-English language background children, the native speakers, and the results for the total sample in terms of Rasch abilities (A_i). An indication of the standard error (SE_{A_i}) is provided for each raw score.

The correlation between the abilities of non-English and English language background children was found to be \( r = 0.98 \), which shows that the assessment provided on the basis of the Rasch scaled _Word Knowledge Test_ is not discriminatory. The assessment of the relationship between the difficulty of an item for a particular individual and his or her ability remains constant from group to group and from item to item. For a further discussion of the assumptions of the Rasch Model of measurement, its advantages, and the fit of the tests contained in the ACER Early School Series, refer to Chapter 10.

**Implications**

Children from non-English-speaking homes and children who for other reasons have a limited background of pre-school experience, who might be shy and withdrawn, who are rarely talked to, or who have inadequate models for language imitation will have an insufficient vocabulary. As has been pointed out previously, language development is vital not only to social adequacy but to intellectual development and school success.

Intervention and remediation procedures need to stress the correct use of words and expressions. Although colloquialisms may be necessary, especially at the beginning, replacement by generally acceptable English usage is desirable. The _Word Knowledge Test_ provides the teacher with an indication of at least some words with which the child is familiar. This assessment may help in the identification of a starting point for vocabulary enrichment. Listening skills need to be developed. Story telling, word and other attention games, practice in following simple and more complicated directions, the use of puppets, etc. have traditionally been found to constitute useful aids for the achievement of this.

The word knowledge of young children is basically comprised of labels for common objects, events, and activities. The general vocabulary of any individual can be enlarged by a wide range of experiences. Material presented in picture books, story telling, the manipulation of common objects, and the discussion of everyday events and experiences have traditionally provided the means to enlarge and consolidate word knowledge. Practice is provided in the naming of objects and parts of objects, and in the analysis of the functional relationships between the parts. The ability of children to communicate their needs and ideas to others has been especially emphasized (Cazden, 1972). Other important skills include the ability to summarize events and to predict outcomes.
Table 9.4 Word Knowledge Test: Conversion of Raw Score to Rasch-scaled Word Knowledge Score (Ai) and Standard Errors (SE_{Ai})

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Total sample</th>
<th>Non-English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ai</td>
<td>SE_{Ai}</td>
<td>Ai</td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>6.1</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>4.4</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>3.6</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>3.2</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>2.9</td>
<td>38</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>2.7</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>41</td>
<td>2.5</td>
<td>41</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>2.4</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>44</td>
<td>2.3</td>
<td>43</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>2.2</td>
<td>44</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
<td>2.2</td>
<td>44</td>
</tr>
<tr>
<td>12</td>
<td>61</td>
<td>2.1</td>
<td>45</td>
</tr>
<tr>
<td>13</td>
<td>47</td>
<td>2.1</td>
<td>54</td>
</tr>
<tr>
<td>14</td>
<td>47</td>
<td>2.0</td>
<td>53</td>
</tr>
<tr>
<td>15</td>
<td>47</td>
<td>2.0</td>
<td>53</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>2.0</td>
<td>52</td>
</tr>
<tr>
<td>17</td>
<td>49</td>
<td>2.0</td>
<td>51</td>
</tr>
<tr>
<td>18</td>
<td>49</td>
<td>2.0</td>
<td>51</td>
</tr>
<tr>
<td>19</td>
<td>50</td>
<td>2.0</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>51</td>
<td>2.0</td>
<td>51</td>
</tr>
<tr>
<td>21</td>
<td>51</td>
<td>2.0</td>
<td>51</td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>2.0</td>
<td>52</td>
</tr>
<tr>
<td>23</td>
<td>53</td>
<td>2.0</td>
<td>53</td>
</tr>
<tr>
<td>24</td>
<td>53</td>
<td>2.0</td>
<td>53</td>
</tr>
<tr>
<td>25</td>
<td>54</td>
<td>2.1</td>
<td>54</td>
</tr>
<tr>
<td>26</td>
<td>55</td>
<td>2.1</td>
<td>55</td>
</tr>
<tr>
<td>27</td>
<td>55</td>
<td>2.2</td>
<td>56</td>
</tr>
<tr>
<td>28</td>
<td>56</td>
<td>2.2</td>
<td>56</td>
</tr>
<tr>
<td>29</td>
<td>57</td>
<td>2.3</td>
<td>57</td>
</tr>
<tr>
<td>30</td>
<td>58</td>
<td>2.4</td>
<td>58</td>
</tr>
<tr>
<td>31</td>
<td>59</td>
<td>2.5</td>
<td>59</td>
</tr>
<tr>
<td>32</td>
<td>60</td>
<td>2.7</td>
<td>60</td>
</tr>
<tr>
<td>33</td>
<td>61</td>
<td>2.9</td>
<td>62</td>
</tr>
<tr>
<td>34</td>
<td>63</td>
<td>3.2</td>
<td>63</td>
</tr>
<tr>
<td>35</td>
<td>65</td>
<td>3.6</td>
<td>65</td>
</tr>
<tr>
<td>36</td>
<td>68</td>
<td>4.4</td>
<td>68</td>
</tr>
<tr>
<td>37</td>
<td>72</td>
<td>6.1</td>
<td>73</td>
</tr>
<tr>
<td>38</td>
<td>72</td>
<td>6.1</td>
<td>73</td>
</tr>
<tr>
<td>\bar{X}</td>
<td>2.8</td>
<td>51.3</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
<td>10.4</td>
<td></td>
</tr>
</tbody>
</table>
It has often been suggested that the size and nature of an individual’s vocabulary not only provides an indication of his or her development and environment, but shows a strong correlation with a person’s general ability (e.g. Binet, 1905; Dunn, 1965; Savage, 1968; Wechsler, 1950, 1958, 1966, 1974). Researchers (e.g. Blank, 1974; Goldstein, 1966; John and Goldstein, 1963; Luria and Yudovich, 1971) have emphasized the importance of the knowledge of words and concepts which have multiple reference, because such knowledge will facilitate categorization, an important cognitive skill. Other areas of importance include the learning of words expressing quantity and quantitative relationships, and the vocabulary reflecting relationships between objects and events in time and space.

Whether or not individualized intervention work has to be provided for a child depends on the teacher’s judgment and assessment. A number of commercially available programs and kits for vocabulary and general language development may provide useful ideas and appropriate activities. An example of such materials is the Peabody Language Development Kit published by American Guidance Service, which is available from the Australian Council for Educational Research. Other sources of practical suggestions include Buchanan (1967), Carroll (1966), Engel (1966), Van Allen and Allen (1969), and many published and unpublished materials listed in the ERIC system.
PART III
FURTHER TECHNICAL INFORMATION
THE MEASUREMENT MODEL
ADOPTED FOR THE ACER EARLY
SCHOOL SERIES

This chapter falls into three parts. A discussion of item difficulty in traditional test theory is followed by a description of the Rasch (1960, 1966a, 1966b) approach to the measurement of test difficulty and person ability. In the third part, an indication is provided of the degree to which the ACER Early School Series tests meet the major assumptions of the Rasch model.

Although this chapter is basically concerned with the discussion of a statistical model (i.e. a model defined by mathematical equations), every attempt has been made to discuss matters in a relatively non-technical fashion, using a minimum of simple algebra to convey essential ideas. For readers who wish to pursue the technical points further, references are provided.

Measurement in Relation to Item Analysis Theory

As was noted in Chapter 3, the interpretation of norm-referenced tests produces an assessment of the child's skills or other characteristics as they relate to those of other individuals. Norm-referenced tests are not designed to provide scores which provide teachers with information concerning what the child is able or unable to do. In the past, most educational tests were norm-referenced, as the major aim of users was to compare, select, and group students for various purposes. The item analysis procedures, traditionally used by test constructors, are appropriate for the purposes of norm-referenced measurement.

Generally speaking, the term 'item analysis' refers to the detailed investigation of the responses of all testees in a trial sample to all test items. Each individual's response to each item on the test is recorded and the statistical characteristics of each item are assessed mathematically.

In traditional item analysis procedures, the two item characteristics derived are item difficulty (or item facility) and a discrimination index. The item difficulty index establishes how difficult an item was for the group who took the test. This index is derived by counting the number of individuals who did not complete the item correctly and by dividing this number by the total number of individuals who attempted the item. Conversely, the facility of an item is provided by the number...
of persons who responded correctly divided by the total number who attempted to solve the item. Theoretically item facility shows how easy the item was for the group of respondents.

In most test manuals and statistical text books, no distinction is made between item difficulty and item facility. Item difficulty is conventionally expressed as the number of persons who passed the item divided by the total number of persons; in other words, in terms of item facility. Item facility is usually expressed as a percentage. As a formula the facility $d$ of an item $i$ is derived as follows:

$$d_i = \frac{n}{N} \times 100$$

where $d_i$ stands for the facility of item $i$.

$n$ stands for the number of persons who completed the item correctly; and

$N$ stands for the total number of persons who attempted the item.

The item discrimination index is a value which shows how well performance on an item separates high achievers from low achievers on the test. A commonly used value is derived from a comparison of the number of individuals who responded correctly to the item and whose performance on the total test places them in the highest 27 per cent of the total group of testees, with the number of individuals who responded correctly to the item, but whose total performance places them in the lowest 27 per cent.

It is obvious that traditional item characteristics obtained in the above summarized manner are strongly dependent on the ability of the particular sample or reference group from which they were obtained. This underlines the importance of the validity of the sampling procedures used in the establishment of norms for norm-referenced tests.

As noted previously, the choice of model on which to base the development of a test has to be determined by the type of use which is to be made of the test results, and on the nature of the variables or the behaviour which are to be assessed.

Traditional item analysis procedures do not provide diagnostic information which describes the child's performance in terms of what he or she can or cannot do. What is required for this type of assessment are item characteristics and indices of ability which are as independent as possible from the particular sample in which they are obtained. Used in the traditional way, normed sets of items cannot provide generalizable diagnostic information concerning the traits measured, because it cannot be assumed that a given item requires a generally definable amount of skill or ability. The proportion of the sample answering the item correctly will increase and decrease as the ability level of the group changes. In the same way the relative ability or skill of an individual assessed by the items will vary according to the ability of the group in which the test is taken.

More than 30 years ago, Gulliksen (1950) suggested:

A significant contribution to item analysis theory would be the discovery of item parameters that remained relatively stable as the item analysis group changed; or the discovery of a law relating the changes in item parameters to changes in the group... As yet there has been no systematic theoretical treatment of measures of item difficulty directed particularly toward determining the nature of their variation with respect to changes in group ability. Neither has the experimental work on item analysis been
Among early attempts to provide a measure of item difficulty which would take into account variations in the ability of samples are those of Thurstone and Tucker. Thurstone's (1947) method of difficulty calibration aimed to compensate for possible changes in group ability level, when two different sets of items given to two separate item analysis groups are to be equated. The following example will describe this method.

A set of items making up Test W is given to group 1 and is then to be equated with a set of items making up Test Q which is given to group 2. A proportion of the items from Test W, which is chosen from all levels of item difficulties contained in the test, is included in Test Q and administered to group 2 as part of Test Q. As a result of this procedure the percentage of respondents who complete these items correctly is established for both groups 1 and 2. Plotting the difficulty values of these items obtained in the two groups against each other (i.e. using group 1 values as the abscissa and the group 2 values as the ordinate) the relationship of the two sets of values is represented graphically. Thurstone's method provides for a smoothed line drawn through these values obtained from items which had been administered to both groups as a basis for translating difficulty values of all items administered to group 2 into difficulty values for group 1.

Unfortunately Thurstone's method does not provide a sample independent index of item difficulty because the item difficulties expressed as a percentage of correct responses must be expected to change with any variation in the ability level of the group as part of which they are obtained.

The measure of item difficulty proposed by Tucker (1952) involves plotting against each raw score group the proportion of individuals in that group who correctly completed a given item. The raw score value at which the item is equally often answered correctly and failed is the difficulty level assigned to the item. There appears to be a lack of research investigating the invariance claimed for this particular index.

An intuitively similar and highly promising approach to item analysis, which attempts to overcome the problems of invariance in indices of item difficulty, is based on the work of the mathematician George Rasch. Rasch (1960) developed a number of mathematical models with the specific intention of developing a means to estimate person parameters (e.g. measures of skills or abilities) and test parameters (i.e. item difficulties) which can be independent of each other.

Despite its undoubted promise (Fischer, 1968, 1974; Keats, 1967, 1971) the model has only recently begun to have an impact on the development of objective tests. In the research literature, the impact of the Rasch model is largely due to the work of Wright (1967), 'one of the first researchers to operationalize the Rasch model' (Whitely and Dawis, 1974, p.163) and his colleagues at the University of Chicago (e.g. Wright and Douglas, 1977; Wright and Mead, 1977; Wright and Panchapakesan, 1969; Wright and Stone, 1979). Fischer (1968, 1974) and his colleagues and doctoral students at the University of Vienna (e.g. Allerup, Kempf, Scheiblechner, Spada, and Sydow) have been responsible for the re-discovery and
development of the Rasch model and its contribution to the psychological and educational literature in Europe.

A major advantage and contribution to test design of the Rasch model is that it, in effect, permits the development of ratio scales of ability estimates and of the difficulty of test items. Tests which are based on a Rasch model of measurement permit interpretative statements such as:

For Lisa, Test A (or item i) is twice as difficult as Test B (or item j). In other words for Lisa the probability of making an error on Test A (or item i) is twice that of the probability of her making an error on Test B (or item j). It might also be shown that Lisa is twice (or several times) as likely to make an error on Test A as John.

Another type of statement which could be made on the basis of ratio scale measurement would be:

* Item a is twice as difficult as item b, therefore the likelihood of failing item a is twice that of failing item b. Mary has done well on item a and similar items. Robert has done well on item b. Mary is half as likely to make an error on either type of item.

Statements such as those presented in the above examples cannot be made in relation to most tests which are based on traditional test theory. For example, it would be incorrect to suggest that a child with a reading age of 11 years reads twice as well or is half as likely to make errors in reading as a child with a reading age of 5 years 6 months. No matter how valid and reliable the intelligence test, it would be nonsensical to suggest that a person with an IQ of 140 is twice as intelligent, or half as likely to make errors, as a person with an IQ of 70.

To summarize, the Rasch (1960, 1961, 1966a, 1966b) model of measurement can lead to tests which, other things equal, can provide estimates of person characteristics (e.g. ability), and item difficulty which are invariant no matter how difficult the test items are, provided that the abilities of the testees are consistent with the difficulties of the test items. In these tests, the ratio of the difficulty of one item (or set of items) to the difficulty of another item (or set of items) is relatively invariant irrespective of the ability level of the persons tested. Ability estimates and estimates of item difficulty based on such a model can certainly be expected to come closer to measurement in the sense of the physical sciences, than would traditional psychometric tests.

Theoretical Assumptions of the Rasch Model

The two parameters of major concern to Rasch as well as to most users of educational tests are:

A = the ability or skill of an individual (or the ability of a group of individuals in the sense that several individuals may perform at the same level of ability or skill),

and

D = the difficulty of an item.

The important assumption made by Rasch, which would seem highly logical to most educators, is that the probability of a person passing an item is a function of the degree of ability \( Am \) of the person to the degree of difficulty \( Di \) of the item (i.e. \( \frac{Am}{Di} \) ) and not of the values of \( Am \) and \( Di \) separately as in traditional models of item analysis.
In other words, a Rasch model assumes the relationship between person ability \( A \) and item difficulty \( D \) to be such that, if in a given test the difficulty \( D_2 \) of Item 2 were greater than the difficulty \( D_1 \) of Item 1, then the probability \( p \) of any person \( m \) in responding correctly to Item 1 should be higher than the probability \( p \) of his or her responding correctly to Item 2.

The converse is also assumed. If the ability estimate \( A_4 \) of Person 4 were higher than the ability estimate \( A_3 \) of Person 3, the probability \( p \) that Person 4 would respond correctly to any given test item \( i \) should be higher than the probability \( p \) that Person 3 would respond correctly to the same item.

Therefore, if the difficulty \( D_2 \) of Item 2 were \( k (k > 0) \) times that of Item 1, and if the ability \( A_4 \) of Person 4 were \( k \) times that of Person 3, then

\[
\frac{A_2}{D_1} = \frac{A_4}{D_2}
\]

no matter what the value of the \( k \) would be.

The theoretical limits of the ratio are 0 to \( \infty \). To provide meaningful results in probability terms, Rasch adopted a mathematical function for the person ability to item difficulty ratio \( \frac{A}{D} \), which increases from 0 to 1 as \( \frac{A}{D} \) increases from 0 to \( \infty \) to provide the probability \( p \) of a person with the ability \( A \) passing an item of difficulty \( D \) in the following way:

\[
\frac{A}{D} = \frac{A}{A + D} = p
\]

Conversely, the probability that this person would not pass an item of difficulty \( D \) is given by

\[
1 - p = 1 - \frac{A}{A + D} = \frac{D}{A + D}
\]

it follows that

\[
\frac{1 - p}{p} = \frac{A}{D}
\]

and

\[
\frac{1 - p}{p} = \frac{D}{A}
\]

In the absence of prior knowledge regarding either the abilities, \( A \), of the individuals making up the sample or the difficulties, \( D \), of the items, Rasch (1960) proposed that the following assumptions be made regarding the ability and difficulty parameters:

1. The best estimate of the ability parameter \( A \) for a particular person \( m \) can be derived from the person's raw score \( r \).
2. The best estimate of the difficulty parameter \( D \) for a given item \( i \) can be derived from the frequency with which the item was correctly completed.

The model makes it possible, on a set of similar items measuring a particular latent trait, to set up an ordered matrix (Keats, 1967; Wright and Stone, 1979) with raw scores as one dimension and items ranked according to frequency of correct responses as the other dimension. Use of this model permits attempts to reproduce, as accurately as possible, the probabilities of passing items in the cells
of an item by 'score group' matrix, where persons who have obtained the same raw score are grouped together. A limitation of the model is that it cannot handle either rows or columns in this matrix which contain a number which is either 0 per cent or 100 per cent of the persons in a raw score group. Persons who receive zero or full scores and items passed by all or none of the persons in the group have thus to be excluded from a Rasch analysis.

For technical details and a step-by-step description of test development procedures utilizing a Rasch model, readers are advised to consult Fischer (1974), Willmott and Fowles (1974), Wright and Stone (1979), and similar texts. The mathematical calculations required for the Rasch calibration of test items are facilitated by the use of logarithmic scales. Based on the exponential function

\[ \text{odds of success} = e^{(\text{ability} - \text{difficulty})} \]

the ability and difficulty estimates obtained are initially in units of a log-odds scale. On this measurement scale, one unit difference between ability and difficulty estimates is expressed as odds of success of \( e : 1 \) (\( e = 2.718 \)).

While a log-odds scale provides a convenient means of representation for theoretical purposes, it is less suitable for practical purposes, because its interpretation requires technical and statistical sophistication which is unnecessary in most contexts in which educational tests are used. For the purposes of presenting Rasch calibrations of difficulty and ability estimates for the ACER Early School Series tests, the log-odds scaled values were therefore mathematically transformed to a more easily interpreted scale which is centered at 50. The procedure utilized was similar to that used by other test developers (Cornish and Wines, 1979; Woodcock, 1973) and is discussed in the literature (Willmott and Fowles, 1974; Woodcock and Dahl, 1971; Wright and Douglas, 1975, Wright and Stone, 1979).

The transformation formula used for the calculation of the difficulty (\( D_i \)) and ability (\( A_i \)) values was as follows:

\[ D_i = d_i \left( \frac{5}{\ln 3} \right) + 50 \]

where \( D_i \) is the transformed difficulty estimate, and \( d_i \) is the difficulty estimate in log-odds units. The transformation factor \( (\frac{5}{\ln 3}) \) (i.e. 4.551) indicates that an increase of 5 units on the transformed scale represents a growth factor of 3 in the log-odds scale. It should be noted that each such scale stands alone. A score of 40 on one scale differs in absolute meaning from a score of 40 on another scale.

For example, a raw score of 6 on the Number Test has an associated ability estimate of 57. Similarly a raw score of 7 on the Verb Tense Test provides an ability estimate of 57. This does not imply that a score of 57 on one scale (Number) is necessarily equivalent to a score of 57 on the other (Verb Tense).

Investigation of the Fit of the ACER Early School Series Tests to the Rasch Model

The possibilities provided by a series of tests of relatively invariant ability and difficulty estimates, and the idea of an instrument which might provide ratio
scaling, were highly attractive for the design of the ACER Early School Series tests.

The purpose of the analyses, which are described in the following section, was to
determine the appropriateness of the Rasch model of item analysis for the tests
contained in the ACER Early School Series.

In the literature, only a few studies concerned with the investigation of the
validity of the model for tests have been reported. Both item and ability estimates
were found to be invariant over different non-random samples investigated by
Anderson, Kearney, and Everett (1968), Brooks (1965), Wright and
Panchapakesan (1969), and Tinsley (1971).

In the present investigation, partly in response to Goldstein’s (1979) reproach
and criticisms, a serious effort was made to test the fit of the assumptions of the
Rasch model by comparing the results obtained on the ACER Early School Series
tests from independent samples with quite different characteristics. The
performance of samples of children from different language backgrounds and of
groups of school beginners of differing maturity and developmental level, as
judged by their teachers, were compared in order to investigate the invariance of
Rasch-based estimates of person ability and item difficulty. Rasch’s assumption
that the ratio of the difficulties between any pair of items would remain constant
across groups was also investigated in relation to the item difficulties obtained in
the ACER Early School Series tests.

Rasch himself investigated the invariance assumption in his tests by drawing
what visually appeared to be the best fitting straight lines. Brooks (1965)
recalculated the best fit of Rasch’s data, using least squares procedures, and found
the latter no more accurate than Rasch’s graphically determined values.

To investigate the claim that, if the model fits the data, the ratio of difficulty of
any given item to any other given item would remain essentially constant,
irrespective of any differences between the groups to whom the test can be
administered, Rasch again used graphical methods. He showed the invariance of
the ratios of the difficulty values in terms of the degree to which the points fell
along a straight line with unit slope, when logarithms of item difficulties were
plotted against the mean logarithms of difficulty for the various ability levels (i.e.
raw score groups) within the same sample of subjects.

A more stringent test of this assumption would, as also suggested by Goldstein
(1979), consist of a comparison of the item difficulty ratios obtained from data
collected in separate and independent samples of persons who differed in ability in
some definable way. This type of empirical check of the model appears not to have
been performed by Rasch, but was attempted in the examination of the model in
relation to the ACER Early School Series.

While checking the fit of the Rasch model to the ACER Early School Series
tests, the study made it possible to explore to some extent the sample-free
assumption of the Rasch item difficulty ratios by comparing data from a priori
different samples.

Specifically the following research questions were explored:

1. Are the Rasch calibrated item difficulties independent of the sample on which
they are based?
<table>
<thead>
<tr>
<th>Table 10.1 Comparison of Rasch-scaled Item Difficulties Calibrated on Different Groups of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Auditory</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
</tr>
<tr>
<td>Recognition of</td>
</tr>
<tr>
<td>Initial Consonant</td>
</tr>
<tr>
<td>Sounds</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>A v. B</td>
</tr>
<tr>
<td>A v. C</td>
</tr>
<tr>
<td>A v. D</td>
</tr>
<tr>
<td>B v. C</td>
</tr>
<tr>
<td>B v. D</td>
</tr>
<tr>
<td>C v. D</td>
</tr>
<tr>
<td>Average correlation = 0.54, SD = 0.15</td>
</tr>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Figure</td>
</tr>
<tr>
<td>Formation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>A v. B</td>
</tr>
<tr>
<td>A v. C</td>
</tr>
<tr>
<td>A v. D</td>
</tr>
<tr>
<td>B v. C</td>
</tr>
<tr>
<td>B v. D</td>
</tr>
<tr>
<td>C v. D</td>
</tr>
<tr>
<td>Average correlation = 0.70, SD = 0.19</td>
</tr>
<tr>
<td>Prepositions</td>
</tr>
<tr>
<td>Verb Tense</td>
</tr>
<tr>
<td>Pronouns</td>
</tr>
<tr>
<td>Negation</td>
</tr>
<tr>
<td>Comprehension</td>
</tr>
<tr>
<td>Word Knowledge</td>
</tr>
</tbody>
</table>
2. Are the ability estimates derived from the Rasch model of measurement independent of the relative ability level of the sample in which they were obtained?

3. Are the ratios of pairs of Rasch-scaled item difficulty values invariant across independent groups of subjects of varying developmental level?

These questions were investigated in relation to all tests included in the ACER Early School Series. In the case of the auditory analysis and conceptual skills tests, the items making up each test were scaled separately for the representative sample and for each of the validation samples. For the language tests, the items were scaled separately for native speakers of English, and for the group consisting of children whose home language was not English, as well as for the total group.

To investigate research question 1, the Rasch-scaled item difficulty values were compared for the different samples. One table in each of the chapters discussing the individual tests of the series contains the Rasch item difficulties and standard errors for each item. No significant differences in the mean difficulties were found between the groups. Question 1 was further investigated by correlating the item scale values. Correlations between the different groups varied somewhat for the different tests, and are shown in Table 10.1.

Within tests the correlation (or average correlation in the case of tests in which more than two independent groups were compared) ranged from an average correlation of \( r = 0.54 \) (SD = 0.15) for the Auditory Discrimination Test to correlations of \( r = 0.98 \) for the Number Test and the Comprehension Test. It should be noted that the samples for the latter two tests were completely independent of one another. In other words, no child in either sample of either test had done both tests.

Over all tests and samples, the average correlation was \( r = 0.77 \), SD = 0.19. This value is substantial, but may appear lower than expected. Inspection of Table 10.1 suggests a number of factors which may have contributed to a lower than expected correlation of the item difficulties obtained from different samples. The tests in which the results of a larger number of samples were compared showed lower average correlations, because of the greater variation in ability between the samples. Other researchers have obtained similar results, and the invariance of Rasch calibrated item difficulty values for extreme groups is still being investigated (Fischer, 1978). Another possible explanation may be that the Auditory Discrimination Test, the Recognition of Consonant Sounds Test and certainly the Word Knowledge Test might be measuring less unitary 'latent traits'.

From a theoretical point of view, it should be noted that it was not Rasch's contention that the difficulty index values per se necessarily remain invariant with changes in group ability level, but rather that the ratios among corresponding pairs of such values remain invariant. This means that members of extreme ability groups may well differ to some degree in their perception of the difficulty of the test, but that the relationship between the items remains the same, irrespective of the ability level of the group tested. The latter concern was investigated under research question 3.

Question 2, concerning the invariance of the Rasch calibrated ability estimates irrespective of the ability level of the sample, can be answered in the affirmative.
Inspection of the relevant tables in each of the test chapters containing the ability estimates and the correlations between the Rasch ability estimates obtained for the different samples for each test, shows a correlation of $r = 1.00$ in the case of most of the tests. Exceptions were again the Auditory Discrimination Test, in which correlations of 0.99, 0.91, and 0.91 were obtained for the comparisons of ability levels obtained for the representative sample and two comparison groups; and the Word Knowledge Test, in which the correlation between the ability estimates for English and non-English language background children was $r = 0.98$.

These correlations are high indeed, and show the stability of the Rasch calibrated ability estimates, in terms of their independence of the sample, and their consistency across all tests of the ACER Early School Series.

Question 3 concerning the invariance of the ratios of Rasch difficulty values for pairs of items across different samples was examined for each test by calculating the ratio between the difficulty values for each pair of items separately for each sample. This was followed by the calculations of t- and F-tests to investigate the differences between the independent samples.

The difficulty ratios for all but the Auditory Discrimination, Recognition of Initial Consonant Sounds, and the Word Knowledge Tests are shown in Tables 10.2 to 10.8 at the end of this chapter. The tables for the former tests have not been reproduced in this Handbook because of their size.

No significant differences were found in the difficulty ratios calculated for different samples in the Recognition of Initial Consonant Sounds, Number, and Figure Formation Tests or in any of the tests of syntactic structures. The t- and F-values are shown in Tables 10.2 to 10.8.

The Auditory Discrimination Test was examined in a slightly different manner. As shown in Table 10.1, the Rasch item difficulties of this test seemed less invariant across the different samples than the difficulty estimates of the items constituting the other tests in the series. The ratios of the difficulties for the Auditory Discrimination Test were thus examined separately for each item; that is, separate F-tests were calculated for each column containing the ratios obtained by relating the difficulty estimate of one item to those of all others.

The F-values for 14 out of the 40 items of the test were significant at or beyond the $p = 0.05$ level of significance. This means that in 35 per cent of the items of the Auditory Discrimination Test, Rasch's assumption of the invariance of the ratios of item difficulty values across samples was not met in the present study.

Anderson, Kearney, and Everett (1968) who investigated the invariance of straight item difficulties, not ratios of difficulty of values obtained by two samples in Rasch-scaled intelligence tests, found that 27 per cent of items failed to fit this assumption of the model at the $p = 0.05$ level. These investigators suggested:

"The 0.05 level of significance is probably too stringent a measure and the 0.10 level is more appropriate when testing the fit of the model. (Anderson et al., 1968, p.237)"

If the $p = 0.10$ level of significance is used in the analyses of the present data, five out of the 40 items, or 12.5 per cent of the items constituting the test, are shown not to support this assumption of the model.

The general characteristics of items which failed to fit Rasch's assumption were
compared with those which conformed to the assumption. No particular pattern was detected which differentiated between the two types of items in so far as their difficulty values, content or position in the test were concerned.

To summarize, it can be said that the examination of the research data collected from a number of independent samples of children on their performances on the tests of the ACER Early School Series showed that the Rasch calibrated item difficulties and the ability estimates derived from the obtained raw scores are independent of the sample on which they are based. Rasch's assumption that the ratios of item difficulty values for pairs of items be invariant across different samples was largely supported by the results of the present investigation.

Finally, the reader may have observed the strong relationship between the difficulty estimates obtained by traditional item analysis procedures and the Rasch calibrated difficulty estimates presented in tables in the chapters describing each test. Over all ten tests, only very few items changed their relative position on the basis of their Rasch-scaled difficulty values, when initial placement had been in terms of their traditional difficulty values. Statistically, this stability of position is to be expected, as the Rasch calibrated estimates of item difficulties are non-linear functions of the traditionally obtained difficulty values.

This fact might give rise to questions concerning the practical usefulness of Rasch item calibration procedures if the resulting products are not essentially different from the more generally known and more easily computed traditional item difficulty values.

The strongest argument for Rasch calibration of tests in which the assumptions of the model are basically satisfied is that Rasch scaling adds two highly desirable properties to the measurements provided by the test:

1. it permits ratio scaling of ability estimates and item difficulties;
2. it avoids to a large extent the major limitation of most psychometric tests (i.e. yielding results which are restricted to samples of particular populations) by providing estimates of person ability and item difficulty which are independent of the sample or population in which they were obtained.

Tables 10.2 to 10.8 present the ratios of item difficulty estimates for paired items of the tests. In each cell, the values of the ratios are represented as proportions in decimals for all samples.
Table 10.2  Number Test: Ratios of Item Difficulty Values for Paired Items, for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.91</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.87</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.85</td>
<td>0.93</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.93</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>0.90</td>
<td>0.94</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.89</td>
<td>0.91</td>
<td>0.98</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.78</td>
<td>0.86</td>
<td>0.90</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.84</td>
<td>0.86</td>
<td>0.92</td>
<td>0.92</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.75</td>
<td>0.81</td>
<td>0.87</td>
<td>0.88</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>0.84</td>
<td>0.90</td>
<td>0.90</td>
<td>0.92</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.67</td>
<td>0.73</td>
<td>0.77</td>
<td>0.77</td>
<td>0.82</td>
<td>0.85</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>0.72</td>
<td>0.77</td>
<td>0.77</td>
<td>0.79</td>
<td>0.83</td>
<td>0.85</td>
</tr>
</tbody>
</table>

1 = 1.01 df 27 NS

Table 10.3  Figure Formation Test: Ratios of Item Difficulty Values for Paired Items, for Five Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.77</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.92</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.23</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.77</td>
<td>0.83</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>0.88</td>
<td>0.88</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.91</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.84</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.93</td>
<td>0.96</td>
</tr>
</tbody>
</table>

F = 1.80 df 4/25 NS
### Table 10.4 Prepositions Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.22</td>
<td>1.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.16</td>
<td>0.95</td>
<td>1.04</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.06</td>
<td>0.87</td>
<td>0.91</td>
<td>1.08</td>
<td>0.97</td>
<td>1.04</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.82</td>
<td>0.86</td>
<td>0.94</td>
<td>1.08</td>
<td>0.97</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.92</td>
<td>0.75</td>
<td>0.79</td>
<td>0.87</td>
<td>0.92</td>
<td>1.02</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>7</td>
<td>0.87</td>
<td>0.71</td>
<td>0.75</td>
<td>0.82</td>
<td>0.87</td>
<td>0.94</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>8</td>
<td>0.92</td>
<td>0.75</td>
<td>0.79</td>
<td>0.87</td>
<td>0.92</td>
<td>1.00</td>
<td>0.89</td>
<td>0.94</td>
</tr>
<tr>
<td>9</td>
<td>0.96</td>
<td>0.78</td>
<td>0.82</td>
<td>0.90</td>
<td>0.96</td>
<td>1.03</td>
<td>1.09</td>
<td>1.03</td>
</tr>
</tbody>
</table>

\[ t = 0.96 \text{ df 35 NS} \]

### Table 10.5 Verb Tense Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.00</td>
<td>0.91</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td>0.87</td>
<td>0.91</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.15</td>
<td>1.15</td>
<td>1.15</td>
<td>1.28</td>
<td>0.82</td>
<td>0.86</td>
<td>0.90</td>
<td>0.94</td>
</tr>
<tr>
<td>6</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td>1.00</td>
<td>0.78</td>
<td>0.84</td>
<td>0.88</td>
<td>0.92</td>
</tr>
<tr>
<td>7</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.92</td>
<td>0.72</td>
<td>0.92</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td>8</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>9</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.75</td>
<td>0.59</td>
<td>0.75</td>
<td>0.63</td>
<td>0.66</td>
</tr>
</tbody>
</table>

\[ t = 1.43 \text{ df 35 NS} \]
## TECHNICAL INFORMATION

### Table 10.6 Pronouns Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.04</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.08</td>
<td>1.04</td>
<td>0.87</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.02</td>
<td>0.97</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.95</td>
<td>0.91</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>0.95</td>
<td>0.91</td>
<td>0.98</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.94</td>
<td>0.90</td>
<td>0.96</td>
<td>0.92</td>
<td>0.94</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>0.95</td>
<td>0.91</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.96</td>
<td>0.92</td>
<td>0.88</td>
<td>0.94</td>
<td>0.96</td>
<td>0.96</td>
<td>1.02</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.92</td>
<td>0.88</td>
<td>0.84</td>
<td>0.90</td>
<td>0.92</td>
<td>0.92</td>
<td>0.98</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.86</td>
<td>0.82</td>
<td>0.78</td>
<td>0.84</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.89</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.79</td>
<td>0.84</td>
<td>0.82</td>
<td>0.77</td>
<td>0.84</td>
<td>0.84</td>
<td>0.87</td>
<td>0.93</td>
<td>0.98</td>
</tr>
</tbody>
</table>

\[ t = 0.94 \text{ df } 54 \text{ NS} \]

### Table 10.7 Negation Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.97</td>
<td>1.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.91</td>
<td>1.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>0.94</td>
<td>0.90</td>
<td>0.90</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.86</td>
<td>0.98</td>
<td>0.94</td>
<td>0.94</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.76</td>
<td>0.87</td>
<td>0.83</td>
<td>0.83</td>
<td>0.92</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.78</td>
<td>0.89</td>
<td>0.85</td>
<td>0.85</td>
<td>0.94</td>
<td>0.90</td>
<td>1.01</td>
</tr>
<tr>
<td>9</td>
<td>0.86</td>
<td>0.67</td>
<td>0.86</td>
<td>0.86</td>
<td>0.82</td>
<td>0.87</td>
<td>0.89</td>
</tr>
</tbody>
</table>

\[ t = 0.07 \text{ df } 27 \text{ NS} \]
Table 10.8 Comprehension Test: Ratios of Item Difficulty Values for Paired Items for Two Independent Samples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.82</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.77</td>
<td>0.92</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.84</td>
<td>0.94</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.78</td>
<td>0.94</td>
<td>0.96</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.84</td>
<td>0.94</td>
<td>1.03</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.74</td>
<td>0.89</td>
<td>0.90</td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td>0.88</td>
<td>0.98</td>
<td>0.94</td>
<td>0.94</td>
</tr>
</tbody>
</table>

\[ t = 1.00 \text{ df 14 NS} \]
Bibliography


BIBLIOGRAPHY


Braine, M.D. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. Psychological Monographs, 1959, 77 (475), whole number.


Calfee, R.C., Kimmoom, P.; and Lindamood, C. Acoustic-phonetic skills and reading in kindergartners to twelfth grade. Journal of Educational Psychology, 1973, 64, 293-8.


Gelman, R. Logical capacity of very young children: Number invariance rules. *Child Development*, 1972, 43, 75-90. (b)


BIBLIOGRAPHY


Rasch, G. An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, 1966, 19, 49–57. (b)


Sanacore, J.A. A checklist for the evaluation of reading readiness. Elementary English, 1974, 50, 858-60.


Smith, M.K. Measurement of the size of general English vocabulary through the elementary grades and high school. General Psychological Monographs, 1941, 24, 311-45.


BIBLIOGRAPHY


APPENDIX I

Glossary of Terms

Since some sections of this Handbook are somewhat technical in nature, certain words may be unfamiliar to the reader and some commonly used words may take on a specialized meaning. In this appendix, a glossary of 154 terms used in educational and psychological measurement is presented to assist readers who encounter an unfamiliar term. The list of meanings presented is not complete. Concepts which have been explained in detail within the body of the Handbook are denoted briefly. Some are not included. For more complete explanations of measurement terms, the reader is referred to Chaplin (1975), Drever (1964), Kendall and Buckland (1975), Wolman (1975), and other texts.

achievement test. A test designed to measure the level of proficiency or the knowledge an individual has attained in some specified area or in general.

age equivalent. The level of an individual's development with respect to any trait or characteristic, that is expressed as equal to the chronological age at which the particular level is normally attained. The age equivalent score expresses the individual's performance as the age at which most individuals reach that particular level of development or performance.

age norms. Indication of age level, usually in years and months, for which a given raw score was the average in the standardization sample.

analysis. The process of reducing a complex phenomenon to simpler and more elementary parts.

answer sheet. Separate test page(s) on which the testee records the selected responses.

anxiety. A state of apprehension or uneasiness related to fear. The object of anxiety is usually less specific than the object of fear (e.g., fear of dogs, spiders, etc.).

aphasia. Impairment or loss of the ability to articulate words or comprehend speech, usually resulting from brain damage.

aptitude test. A set of tasks chosen and standardized so as to allow for a prediction of an individual's future performance on tasks which are similar to the tasks on the test.

attention. The focusing of perception leading to heightened awareness of a limited range of stimuli.

attention span. (a) The duration of time a person can concentrate on one event; (b) the amount of material or the number of separate items which an individual can note during one brief exposure.

autism. An extreme inability to relate to other people resulting from an absorption in (often bizarre) fantasy to the exclusion of interest in reality; it may be linked with organic causes or with schizophrenia.
autistic child. A child who displays autistic symptoms such as extreme withdrawal, complete lack of interest in people, language disturbance, often mutism, severe apathy, and emotional detachment. Other behaviours frequently observed in autistic children include repetitive head banging, rocking etc., perseveration, and extreme fear of and resistance to change.

battery of tests. A set of tests intended to be administered successively, and the composite scores of which are used to assess individual differences.

bimodal distribution. A distribution in which frequencies tend to concentrate or pile up at two distinct points or regions.

calibration (of tests or test items). The process of translating the scores obtained on an instrument into an expression based on a known external or internal standard.

central tendency. An average value of a set of scores, i.e. mean, median, or mode.

comparable scores. Expressed on the same scale and with the same relative meaning within some common reference group. If scores on different tests are comparable, a particular numerical score represents the same level of proficiency or deficiency regardless of the subject matter of the test.

concept. The properties or relationships common to a class of phenomena or ideas. Concepts may be concrete or abstract.

calibration (of tests or test items). The process of translating the scores obtained on an instrument into an expression based on a known external or internal standard.

central tendency. An average value of a set of scores, i.e. mean, median, or mode.

comparable scores. Expressed on the same scale and with the same relative meaning within some common reference group. If scores on different tests are comparable, a particular numerical score represents the same level of proficiency or deficiency regardless of the subject matter of the test.

concept. The properties or relationships common to a class of phenomena or ideas. Concepts may be concrete or abstract.

central tendency. An average value of a set of scores, i.e. mean, median, or mode.

comparable scores. Expressed on the same scale and with the same relative meaning within some common reference group. If scores on different tests are comparable, a particular numerical score represents the same level of proficiency or deficiency regardless of the subject matter of the test.

concept. The properties or relationships common to a class of phenomena or ideas. Concepts may be concrete or abstract.

central tendency. An average value of a set of scores, i.e. mean, median, or mode.

comparable scores. Expressed on the same scale and with the same relative meaning within some common reference group. If scores on different tests are comparable, a particular numerical score represents the same level of proficiency or deficiency regardless of the subject matter of the test.

concept. The properties or relationships common to a class of phenomena or ideas. Concepts may be concrete or abstract.

central tendency. An average value of a set of scores, i.e. mean, median, or mode.

comparable scores. Expressed on the same scale and with the same relative meaning within some common reference group. If scores on different tests are comparable, a particular numerical score represents the same level of proficiency or deficiency regardless of the subject matter of the test.

concept. The properties or relationships common to a class of phenomena or ideas. Concepts may be concrete or abstract.
**dependent variable.** A variable in which changes are attributed to changes in the independent variable. In experiments the dependent variable is often a response to a stimulus variable.

**deviation.** The amount by which a score differs from some reference measure of the distribution of scores, e.g. the numerical difference \((X - \bar{X})\) of a given raw score \(X\) from the mean \(\bar{X}\) of the distribution.

**diagnostic test.** A test designed to identify the nature and source of the individual's strengths or difficulties.

**difficulty.** The degree of arduousness, represented by the difficulty index \((D_i)\), of a test or item for individuals, as opposed to its easiness or facility.

**difficulty index \((D_i)\).** Derived from the proportion of testees in a group who answered the item incorrectly.

**discrimination.** Degree to which performance on an item separates students with high and low scores on the total test.

**discrimination index.** A value which indicates the ability of an item to separate high-achieving students from low-achieving students. A commonly used index is obtained by comparing the number in the highest 27 per cent of the group who responded correctly with the number in the lowest 27 per cent who responded correctly.

**distractor.** Any of the incorrect response options in a multiple-choice test item.

**distribution of scores.** Tabulation or enumeration of the frequency of occurrence of each score in a given set of scores.

**domain-referenced test.** Test aimed to assess knowledge or skills pertaining to a completely specified area, situation, or principle.

**dyslexia.** Severe difficulty or inability in learning to read. A specific learning difficulty.

**easiness.** See facility.

**empathy.** Ability to perceive the mood and feelings of another person. Understanding of feelings, sufferings, etc. of another person which have not been communicated in words.

**empirical.** Based on facts and experience, systematic observation, and experiment rather than on theory or general philosophical principles.

**equivalent forms.** The items of equivalent forms of a test are of the same type, cover the same content, have the same distribution of difficulty values, and result in scores of the same mean, variability, validity, and reliability.

**error of measurement.** The difference between an obtained score and the corresponding hypothetical true score.

**error variance.** In a set of test scores, the term refers to the mean of the square errors of measurement for each score in the set.

**etiology.** The investigation of the causes or significant antecedents of a given phenomenon.

**evaluation.** Judgment of merit, sometimes based solely on measurements such as those provided by test scores, but more frequently involving a synthesis of various measurements, critical incidents, subjective impressions, and other kinds of evidence.

**face validity.** A measure of the degree to which the items in a test appear, to tester and testee, to be measuring the characteristics which the test is intended to measure.

**facility index.** A measure of the percentage of correct responses determined by dividing the number of students getting the item right by the number who tried the item. Used to establish how easy an item was for the group who took the test. See also difficulty index.

**factor.** Underlying influence accounting for part of the variability of a number of scores or other behavioural manifestations.

**factor analysis.** Statistical procedure which seeks to identify a small number of hypothetical constructs which might account for the intercorrelations between scores on a much larger number of tests or other phenomena. The factors resulting from factor analytic procedures are weighted sums of observed variables which are expected to summarize, describe, and explain the observed data.

**fit (statistical).** The level of agreement of obtained data to a predetermined standard, or the adjustment of data to such a standard.
frequency distribution. An ordered sequence of score intervals, which contains the frequencies with which scores falling into the specified intervals occurred in the sample.

grade equivalent score. Report of a student's performance on a test in terms of the school grade level for which that level or performance is typical.

homogeneity. Similarity found among members of a set of persons, objects, or data. More specifically the degree to which the items comprising a test measure a single variable or construct.

hypothetical construct. Concept going beyond observation derived from a framework of theory with the aim of explaining behavioral data (e.g., intelligence, learning, motivation).

IQ (intelligence quotient). Originally a ratio of the individual's mental age to his or her chronological age. On some more recent intelligence tests, IQ may be a standard score most frequently from a distribution with a mean of 100 and a standard deviation of 15.

independent variable. The variable under experimental control in relation to which changes under observation are studied. In experiments, the independent variable is often a stimulus, responses to which are the dependent variables.

intelligence. The capacity to understand and manage facts and propositions, relationships and symbols. The ability to reason.

intercorrelation. A term used to refer to the correlation of a number of variables among themselves, as distinct from the correlations between them and an 'external' or dependent variable.

interval scale. A type of measurement scale which allows not only the ordering by magnitude of scores reflecting the characteristic being measured, but which, though not possessing an absolute zero point, possesses equal intervals of magnitude.

invariance. Remaining constant despite changes in other conditions.

item analysis. An examination of student performance for each item on a test. It consists of re-examination of the responses to items of a test by applying mathematical techniques to assess two characteristics—difficulty and discrimination—of each objective item on the test.

item-test correlation. Correlation between the scores obtained on an item and the scores obtained on the test as a whole.

key. A solution set containing the correct responses to objective test items.

latent trait. This term refers to a trait or variable which cannot be observed directly but which is assumed to enter into the structure of a system or phenomenon under study, e.g., intelligence, creativity, anxiety; etc.

lateral dominance. The tendency for dominance of one side of the brain over the other in most functions. More specifically, preferential use of one hand or side of the body.

mastery test. A test not intended to indicate how much a student has achieved relative to other students, but only whether or not he or she has achieved enough to satisfy the minimum requirements of the teacher or the examiner.

matrix. A table of numbers arranged in rows and columns which is subjected to certain mathematical procedures.

mean. A measure of central tendency. The mean provides an average numerical value of a set of scores. The arithmetic mean is calculated by adding all scores and by dividing the sum by the number of scores summed.

measurement. A process of assigning numbers to the individual members of a set of objects or persons for the purpose of indicating differences among them in the degree to which they possess the characteristic being measured.

median. A measure of central tendency which represents the point in a distribution of scores, which divides the distribution into two parts, such that an equal number of scores fall above and below the point, i.e., the 50th percentile or midpoint of the distribution.
mental age. The score on a test of mental ability expressed in terms of the chronological age of persons whose average test score is the same as his or hers.

mode. Measure of central tendency which indicates the most frequently occurring score or value in a frequency distribution.

morpheme. The smallest unit in language (in a word, word-stem, prefix, ending) which is still meaningful. See also phoneme.

multiple-choice item. An item consisting of a stem, i.e. a direct question, direction, or incomplete statement, and two or more options from which to choose the answer or correct completion of the item.

nominal scale. The weakest type of measurement scale. It labels objects or observations (usually by letter or number) for identification or classification purposes. This type of scale does not reflect magnitude relationships, and does not possess equal intervals or an absolute zero point.

norm. The value reflecting the performances of a defined group on a test.

norm-group. Group of persons whose performance on a test provides a standard of comparison for subsequent test takers.

norm-referenced measurement. Assessment based on norm-referenced tests.

norm-referenced tests. Tests which are designed to describe the testee's status or achievement in relation to the performance of a group of other testees. Usually the latter group will have completed the test at some earlier stage, and is described as the norm group. Because the performance of subsequent groups is referenced to or compared with the norm group, the term 'norm-referenced' is used.

normal distribution. A mathematically defined, idealized frequency distribution. It is represented by a symmetrical bell-shaped curve characterized by scores concentrated around the average and decreasing and tapering off towards each extreme.

normalized standard scores. A transformation of scores that results in an approximately normal distribution regardless of the shape of their initial distribution.

norms. These provide the basis for the interpretation of scores on standardized tests by showing how the members of a particular reference group or groups performed on the test.

objectivity. A characteristic of statements that can be verified by independent observation or judgment.

operational definitions. These describe how the variable being so defined can be quantified, i.e. how amounts of the variable can be determined.

ordinal scale. A type of measurement scale which arranges scores with reference to their magnitude and assigns numbers accordingly, i.e. first, second, third. Ordinal scales lack an absolute zero point, and are not made up of equal intervals.

Pearson's 'r'. See product moment correlation.

percentile rank. Number indicating the percentage of scores in the whole distribution falling below the point at which a given score lies.

performance test (or item). A test (or item) which requires the testee to demonstrate his or her skill or ability by manipulating objects or instruments. Such a test minimizes the role of language.

phoneme. Similar speech sounds which serve to distinguish utterances from one another.

phonetics. A systematic representation of speech sounds by means of symbols.

point-biserial correlation. A correlation between a dichotomous and a continuous variable.

population. The total set of persons, items, or observations pertaining to a particular universe. Some part of the population becomes a sample.

power test. A test which is designed to measure the level of achievement or ability a person can reach usually without regard to speech.

predictive validity. Indicates how accurately performance on a test predicts performance on some later measure or outside criterion.

product moment correlation coefficient (or Pearson's 'r'). Obtained by calculating the mean of the products of the standard scores of the two measures being correlated across all individuals in the group.
profile. Graphic representation of the relative magnitude of an individual's scores on several tests. The procedure is used to represent the individual's skills and development by establishing a pattern based on his or her performance on a number of tests yielding comparable scores.

psychometrics. Specialized branch of psychology which deals with individual differences and their assessment by means of psychological tests.

psychomotor. The area of human action which emphasizes all types of body movements which are involuntary or voluntary.

case. A number of cases of any sort drawn from a population or total set in such a way that every item in the population or set has an equal chance of being chosen and that the choice of each item is independent of that of any other.

range of scores. The smallest-interval on the scale of scores which includes all scores in the distribution. Statistically the range is a measure of variability, and its size is computed by subtracting the lowest score in a distribution from the highest score.

rank. The position of an item or other variable in relation to others which have been arranged according to some specific criterion.

rank order. The arrangement of a series of values, scores, individuals, observations, etc. in order of (increasing or decreasing) magnitude or size. The intervals between ranked values are not necessarily equal.

rapport. The relationship which exists between tester and testee.

ratio. A relationship between two phenomena variables, etc. in number, degree, or quantity. More specifically, a quotient, as in intelligence quotient (IQ), which is equal to the product of mental age (MA) divided by chronological age of the testee (CA), i.e. IQ = MA/CA.

ratio scale. A type of measurement scale which reflects the relative magnitude of scores, is based on an absolute zero point, and reflects equal intervals which can be added and divided.

rationale. The basic or underlying reason or justification for a hypothesis, opinion, or action.

raw score. The score as originally obtained on the test, before any transformation to a standard score or other derived score has been calculated.

reading readiness. Denotes the stage at which the child because of developmental, maturation, and situational factors is able to profit by certain conditions of reading instruction.

reasoning. Rational and logical thinking.

referral. Sending a client to another professional agency or institution.

reliability. The consistency with which the test measures whatever it measures.

representative sample. A sample chosen in such a way as to make it more likely to exhibit the same characteristics as the population. A random sample is a representative sample.

response set. A predisposition or tendency on the part of the testee to answer a test item in a particular way. Willingness or unwillingness to guess may be a response set.

restriction-of-range. See truncation.

rote learning. Memorization of a sequence of words or other symbols by repeated utterance or observation. Rote learning does not rely on meaningfulness as an aid to learning.

sample. Any subset of persons, items, or observations which are selected to represent a larger group or set of observations.

sampling error. The difference between the sample value of some statistic and the value obtained when calculated on the basis of the entire population.

scaling. Converting raw data into types of scores which are more readily interpreted, e.g. ranks, standard scores.

scatter diagrams. A graphical device used to display the relationship between scores on two tests for the individuals in a group. Scores on one test are represented on the vertical dimension, those on the other along the horizontal dimension. A tally mark is entered on the diagram to reflect the pair of scores for a particular individual.
schizophrenia. A psychotic disorder characterized by fundamental disturbances in the relationship to reality, formation of concepts, and varying degrees of behavioural, affective, and intellectual disturbances. Symptoms may include hallucinations and delusions, and various autistic behaviours.

score. Number assigned to an examinee to provide a quantitative description of his or her performance on a particular test.

scoring key. Indicates the correct answer to each item.

skewed distribution. An asymmetrical distribution in which most of the scores are closer to one end of the distribution than they are to the other. If the longer tail of the distribution extends towards the lower end of the score scale, the distribution is said to be negatively skewed. If the longer tail extends to the higher end of the score scale, the distribution is said to be positively skewed.

speed test. The primary emphasis in this type of test is on the number of items which the testee completes correctly in a given time.

speededness (of a test). The extent to which a testee's score on the test depends on quickness in working through it. It is sometimes measured by the proportion of examinees who do not reach and answer the last item in the test.

split-halves reliability coefficient. A coefficient which is obtained by using half the items on the test, sometimes the odd-numbered items, to yield one score for an examinee and the other half of the items to yield another independent score. The correlation between the scores on these two half-tests, corrected with the aid of the Spearman-Brown Formula, provides an estimate of the reliability of the total test.

standard. Any measure, rule, or principle which is used as a basis for comparison or judgment.

standard deviation. The most widely used measure of dispersion of a frequency distribution. It provides an indication of the spread of scores about the mean of the distribution, and is equal to the positive square root of the variance.

standard error. An estimate of the standard deviation of the errors associated with a variable or test scores in a given sample. Mathematically it is the positive square root of the variance of the sampling distribution of a statistic.

standard error of measurement. An estimate of the (averaged) standard deviations of errors associated with each score.

standard score. A score derived from a raw score so that it can be expressed on a uniform standard scale without seriously altering its relationship to other scores in the distribution. A simple type of standard score is the z-score, which expresses each raw score as a positive or negative deviation from the mean of all raw scores on a scale in which the unit is one standard deviation.

standardization (of tests). The establishment of norms for the interpretation of scores by giving a test in a standard way to a representative sample or population, and by making studies of the reliability and validity of the test.

stanine. Abbreviation from 'standard nine', a standard score based on a nine-unit scale. The distribution of stanine scores in the population from which they were derived has a mean of 5 and standard deviation of 2.

statistic. A summary value calculated from a sample of observations usually but not necessarily an estimator of some population parameter; a function of sample values.

statistics. (a) Numerical data relating to a set of individuals or observations; (b) the mathematics of collecting, analysing, summarizing, and interpreting such data.

sum of square: \( \sum (X - \bar{X})^2 \). The total of the squares of the deviations of scores from the mean.

syntax. A term derived from a Greek word meaning to order or arrange. As applied to language, it generally refers to the entire network of rules which underlie the structural organization of language. In the narrower sense, the meaning of syntax is restricted to those organizational rules which define the relationships of words in sentences. Syntax can be said to deal with the structure of sentences; sometimes the term is used synonymously with 'grammar'.

synthesis. Combining elements or components into a whole.
test. This term is used here in its most common meaning, as an instrument or other means of assessment which is standard in format.

test relevance. Determined by the extent to which the test contributes to the aims of the tester.

test-taking skill. Knowledge and experience of taking a test which assists the test taker to use test time wisely and to respond appropriately to items, as well as coping with other elements of the testing session.

testee. A person to whom a test is being administered.

text. A set of assumptions advanced to explain phenomena in relation to existing data, and used to predict new events or instances.

trait. An inherited or acquired characteristic or set of characteristics which is consistent and stable.

true score. An idealized error-free score for an individual on a specified test. The value of an observation entirely free from error. Also defined as the mean of an infinite number of scores obtained in repeated performances on the test.

truncation. A truncated distribution results if part of a total distribution, lying on either side of a fixed value, is ignored or cut off.

validation. The process of determining whether a test measures what it was designed to measure.

variable. Generally, any quantity which varies. In the mathematical sense, it is a quantity which can take on one of a specified set of values. The term is also applied to non-quantifiable variables, e.g. 'sex' is a variable in this sense since any child may be described by one of two 'values', boy or girl.

variance. A measure of the dispersion or variability of the scores in a distribution. Mathematically, the variance is the mean of the squared deviations of the scores from their mean.

weight. The importance of an object in relation to a set of objects to which it belongs. A numerical coefficient attached to an observation, frequently by multiplication in order that it will assume a derived degree of importance in a function of all the observations of the set. In 'weighted scoring', the number of points awarded for a correct response is not the same for all items in the test. In some cases, weighted scoring involves the award of different numbers of points for the choice of different responses to the same item.
**APPENDIX II**

Table A.1 Point-biserial Correlations and KR 20 Reliability Coefficients for All Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Item</th>
<th>r\text{bis}</th>
<th>Item</th>
<th>r\text{bis}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Discrimination</td>
<td>1</td>
<td>0.54</td>
<td>21</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.54</td>
<td>22</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.54</td>
<td>23</td>
<td>0.48</td>
</tr>
<tr>
<td>KR 20 = 0.92</td>
<td>4</td>
<td>0.54</td>
<td>24</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.68</td>
<td>25</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.68</td>
<td>26</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.68</td>
<td>27</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.68</td>
<td>28</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.73</td>
<td>29</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.57</td>
<td>30</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>0.49</td>
<td>31</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.42</td>
<td>32</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>0.29</td>
<td>33</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.29</td>
<td>34</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0.59</td>
<td>35</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0.61</td>
<td>36</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>0.54</td>
<td>37</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>0.44</td>
<td>38</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>0.59</td>
<td>39</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.55</td>
<td>40</td>
<td>0.45</td>
</tr>
<tr>
<td>Recognition of Initial Consonant Sounds</td>
<td>1</td>
<td>0.74</td>
<td>10</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.77</td>
<td>11</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.80</td>
<td>12</td>
<td>0.76</td>
</tr>
<tr>
<td>KR 20 = 0.94</td>
<td>4</td>
<td>0.75</td>
<td>13</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.71</td>
<td>14</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.70</td>
<td>15</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.67</td>
<td>16</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.72</td>
<td>17</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>1</td>
<td>0.47</td>
<td>5</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.53</td>
<td>7</td>
<td>0.63</td>
</tr>
<tr>
<td>KR 20 = 0.70</td>
<td>3</td>
<td>0.65</td>
<td>7</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.59</td>
<td>8</td>
<td>0.55</td>
</tr>
</tbody>
</table>

133

140
<table>
<thead>
<tr>
<th>Test</th>
<th>Item</th>
<th>( r_{bis} )</th>
<th>Item</th>
<th>( r_{bis} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure Formation</td>
<td>1</td>
<td>1.00</td>
<td>6</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.72</td>
<td>7</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.73</td>
<td>8</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.67</td>
<td>9</td>
<td>0.59</td>
</tr>
<tr>
<td>Prepositions</td>
<td>1</td>
<td>1.00</td>
<td>6</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.39</td>
<td>7</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.46</td>
<td>8</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.67</td>
<td>9</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb Tense</td>
<td>1</td>
<td>0.34</td>
<td>6</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.52</td>
<td>7</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.55</td>
<td>8</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.62</td>
<td>9</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pronouns</td>
<td>1</td>
<td>0.68</td>
<td>7</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.54</td>
<td>8</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.50</td>
<td>9</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.54</td>
<td>10</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.68</td>
<td>11</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negation</td>
<td>1</td>
<td>0.53</td>
<td>5</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.63</td>
<td>6</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.50</td>
<td>7</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.55</td>
<td>8</td>
<td>0.63</td>
</tr>
<tr>
<td>Comprehension</td>
<td>1</td>
<td>0.28</td>
<td>4</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.41</td>
<td>5</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.57</td>
<td>6</td>
<td>0.64</td>
</tr>
</tbody>
</table>
General Suggestions Concerning Intervention and Remediation

The ACER Early School Series consists of 10 tests which provide teachers with a small sample of the child's functioning in major areas of cognitive and language development. The test scores will indicate to the teacher whether a particular child's performance on the skills assessed is developmentally within the range of expectation, is advanced, or is lower than might be expected. In addition to this, the performance scores provide the teacher with information as to the kinds of experiences and teaching which can be expected to be most beneficial for each child.

The ideas and suggestions for intervention provided in the intervention sections of Chapters 4 to 9 are not intended to be prescriptive. They are suggestive only of approaches and activities which might help towards the development and achievement of particular skills. The same is true for the general points noted in this appendix, the aim of which is to bring certain general points, which have been found useful by teachers, to the attention of the users of the ACER Early School Series.

A restatement of the following general suggestions concerning the design and implementation of intervention and remedial procedures, though familiar to most teachers, might be useful:

- Above all, make learning fun.
- Steps to success will be achieved in appropriately selected, small, easy, sequential learning exercises.
- Guard against (where necessary, help to overcome) feelings of failure in the child.
- Start intervention some steps below the functioning level of the individual.
- Select materials and techniques which have not previously been tried with this individual.
- Use a lot of encouragement and praise. (Find out what type of encouragement is reinforcing to the individual.)
- Utilize as much concrete material as possible.
- In the initial stages, utilize the preferred or strongest sense modality. Then gradually change over to the weaker modality.
In the initial stages of remediation, provide as much individual attention as possible. (Best results would be expected in one-to-one situations.)

An active learner learns faster than a passive one. Activities should include moving, touching, manipulating, feeling, trial and error activity, as well as hearing and seeing. It is especially important for the individual to be involved in activities and games which relate directly to weaker areas.

Most teachers accept the validity of a variety of approaches to intervention and remediation, and advocate that the method of instruction should be adjusted to the skills and needs of individual learners. An important source of individual differences and developmental needs among school beginners lies in the fact that some children show marked discrepancies in skills and their preferences between visual and auditory perception. Studies such as de Hirsch, Jansky and Langford (1966), Suchman and Trabasso (1966), and Dechant (1973) have provided an indication of wide differences in modality preference and their possible influence on learning. The importance of the relationship between modality preference-based methods of early reading instruction and reading achievement has been stressed in the research literature (e.g. Bateman, 1968; Bruininks, 1969; Daniel and Tacker, 1974; Donovan and Austin, 1978; Robinson, 1972; Wepman and Morency, 1971; Wolpert, 1971).

Although stimulus perception is only one component of the learning process, research and experience have shown that strong modality preferences should be taken into consideration in the choice of teaching methods for school beginners. For pupils who may learn more easily through the preferred sense modality, or who may find themselves handicapped in teaching methods emphasizing one of the modalities, the most appropriate method should be identified. It should be remembered that, even in individuals with extreme modality preferences, an appropriate match between preference and method does not ensure that learning takes place.

Most teachers observe modality preferences among pupils early in the school year. This is of particular importance in the case of physically handicapped children.

<table>
<thead>
<tr>
<th>Modality preference</th>
<th>Auditory</th>
<th>Visual</th>
<th>No preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory preference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No preference</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where strong modality preferences or physical disability of a modality are suspected, the Illinois Test of Psycholinguistic Abilities, ITPA (Kirk et al., 1968) provides a useful means of further information concerning modality preferences of an individual. Other tests which have been found useful in follow-up investigations of children with strong modality preferences include the Gates MacGinitie Readiness Tests, GMRT, (Gates and MacGinitie, 1965).

On the basis of the profiles of scores obtained on these tests, modality preferences can be defined as follows:

*Auditory preference* is shown when the ITPA sequential memory sub-test scaled score is nine or more points higher than the visual sequential memory sub-test score, and/or when the average stanine score on the four auditory tests of the
GMRT is two stanines higher than the average stanine score on the four visual tests of the GMRT.

Visual preference is established when the ITPA visual sequential memory score is higher than the individual's scaled score on the auditory memory test and/or the obtained average stanine score on the four visual tests of the GMRT is two stanines higher than the average stanine on the four auditory tests of the GMRT.

No preference might be defined by an obtained scaled score difference of ITPA visual sequential memory and auditory sequential memory tests of less than nine points and/or an obtained difference of zero or one mean stanines between the four auditory and the four visual tests of the GMRT.

Appropriate methods of reading instruction for children with strong auditory modality preference would emphasize phonics at least in the initial stages. A number of 'structural-reading programs' are based on the analysis of phonics and stress sound/symbol associations. An example of this type of program is the Structural Reading Program (Stern, Gould, and Stern, 1972).

A program emphasizing the visual modality is The Hawaii English Program (an individualized, multi-media, self-instructional or peer-taught program with three basic content areas: language skills, language systems, and literature). Direct teaching of sound/symbol associations and other phonic analysis cues are not part of this instructional program. Instead, success in this program depends on the learner's visual discrimination skills and usage of visual cues and visual memory.

For children without marked modality preferences, an eclectic approach of instructional procedures may afford the strongest probability of success. Such a program emphasizes word recognition and comprehension. Structural analysis, phonics, contextual cues, configurational and linguistic pattern drill, and dictionary skills are stressed. There is a tendency to utilize visual and auditory channels equally and, to some degree, simultaneously. Examples of such programs include Breakthrough to Literacy and Reading 360.
APPENDIX IV

Table A.2  Word Knowledge Test: Source List of Words

<table>
<thead>
<tr>
<th>Abie</th>
<th>Aseptic</th>
<th>Big</th>
<th>Burn</th>
</tr>
</thead>
<tbody>
<tr>
<td>About</td>
<td>Ask</td>
<td>Bigger</td>
<td>Bus</td>
</tr>
<tr>
<td>Accident</td>
<td>Asleep</td>
<td>Bike</td>
<td>Bush</td>
</tr>
<tr>
<td>Across</td>
<td>Assaulting</td>
<td>Bird</td>
<td>But</td>
</tr>
<tr>
<td>Active</td>
<td>At</td>
<td>Birthday</td>
<td>Butterfly</td>
</tr>
<tr>
<td>Adventure</td>
<td>Attack</td>
<td>Biscuit</td>
<td>Button</td>
</tr>
<tr>
<td>Aeroplane</td>
<td>Away</td>
<td>Bit</td>
<td>Buy</td>
</tr>
<tr>
<td>Affliction</td>
<td>Awoke</td>
<td>Bite</td>
<td></td>
</tr>
<tr>
<td>Afraid</td>
<td>Baby</td>
<td>Black</td>
<td>Caboose</td>
</tr>
<tr>
<td>After</td>
<td>Back</td>
<td>Blame</td>
<td>Cake</td>
</tr>
<tr>
<td>Afternoon</td>
<td>Bag</td>
<td>Bless</td>
<td>Calendar</td>
</tr>
<tr>
<td>Afterwards</td>
<td>Balance</td>
<td>Block</td>
<td>Call</td>
</tr>
<tr>
<td>Again</td>
<td>Ball</td>
<td>Body</td>
<td>Came</td>
</tr>
<tr>
<td>Age</td>
<td>Bar</td>
<td>Book</td>
<td>Camp</td>
</tr>
<tr>
<td>Ago</td>
<td>Bark</td>
<td>Both</td>
<td>Can</td>
</tr>
<tr>
<td>Air</td>
<td>Base</td>
<td>Bottle</td>
<td>Cannot</td>
</tr>
<tr>
<td>All</td>
<td>Basket</td>
<td>Bottom</td>
<td>Can't</td>
</tr>
<tr>
<td>Allow</td>
<td>Bat</td>
<td>Bought</td>
<td>Capsule</td>
</tr>
<tr>
<td>All_right</td>
<td>Bath</td>
<td>Bow</td>
<td>Car</td>
</tr>
<tr>
<td>Alone</td>
<td>Be</td>
<td>Bowl</td>
<td>Care</td>
</tr>
<tr>
<td>Along</td>
<td>Beach</td>
<td>Boy</td>
<td>Carry</td>
</tr>
<tr>
<td>Alphabet</td>
<td>Bear</td>
<td>Branch</td>
<td>Cart</td>
</tr>
<tr>
<td>Am</td>
<td>Became</td>
<td>Brave</td>
<td>Case</td>
</tr>
<tr>
<td>Angry</td>
<td>Because</td>
<td>Breed</td>
<td>Catch</td>
</tr>
<tr>
<td>Animal</td>
<td>Bed</td>
<td>Break</td>
<td>Catcher</td>
</tr>
<tr>
<td>Another</td>
<td>Bedroom</td>
<td>Brick</td>
<td>Caught</td>
</tr>
<tr>
<td>Answer</td>
<td>Bee</td>
<td>Bridge</td>
<td>Chair</td>
</tr>
<tr>
<td>Ant</td>
<td>Been</td>
<td>Bring</td>
<td>Change</td>
</tr>
<tr>
<td>Any</td>
<td>Beetroot</td>
<td>Broke</td>
<td>Chase</td>
</tr>
<tr>
<td>Anyone</td>
<td>Before</td>
<td>Break</td>
<td>Chattel</td>
</tr>
<tr>
<td>Anything</td>
<td>Beg</td>
<td>Break</td>
<td>Chemist</td>
</tr>
<tr>
<td>Apple</td>
<td>Behind</td>
<td>Brick</td>
<td>Chicken</td>
</tr>
<tr>
<td>Appliance</td>
<td>Bell</td>
<td>Bridge</td>
<td>Child</td>
</tr>
<tr>
<td>April</td>
<td>Best</td>
<td>Bridge</td>
<td>Chimney</td>
</tr>
<tr>
<td>Aquarium</td>
<td>Better</td>
<td>Brown</td>
<td>Chisel</td>
</tr>
<tr>
<td>Archer</td>
<td>Bewail</td>
<td>Brown</td>
<td>Christmas</td>
</tr>
<tr>
<td>Arctic</td>
<td>Bell</td>
<td>Buck</td>
<td>Clean</td>
</tr>
<tr>
<td>Arm</td>
<td>Best</td>
<td>Bubble</td>
<td>Clear</td>
</tr>
<tr>
<td>Around</td>
<td>Better</td>
<td>Build</td>
<td>Climb</td>
</tr>
<tr>
<td>Alpine</td>
<td>Bewail</td>
<td>Build</td>
<td>Clock</td>
</tr>
<tr>
<td>As</td>
<td>Bicycle</td>
<td>Bump</td>
<td>Climb</td>
</tr>
<tr>
<td>As</td>
<td></td>
<td>Bunny</td>
<td>Clock</td>
</tr>
</tbody>
</table>

138
<table>
<thead>
<tr>
<th>SOURCE LIST OF WORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>G</td>
</tr>
<tr>
<td>H</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>J</td>
</tr>
<tr>
<td>K</td>
</tr>
<tr>
<td>L</td>
</tr>
<tr>
<td>M</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>O</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>Q</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>S</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>U</td>
</tr>
<tr>
<td>V</td>
</tr>
<tr>
<td>W</td>
</tr>
<tr>
<td>X</td>
</tr>
<tr>
<td>Y</td>
</tr>
<tr>
<td>Z</td>
</tr>
</tbody>
</table>

**Example entries:**
- `A`: ant, am, ao, A
- `B`: ble, bi, bo, B
- `C`: cde, ce, co, C
- `D`: def, de, do, D
- `E`: efg, ef, et, E
- `F`: fgh, fi, fo, F
- `G`: ghi, gk, gl, G
- `H`: hijk, hj, hl, H
- `I`: im, in, io, I
- `J`: jk, jl, jm, J
- `K`: kn, kl, km, K
- `L`: ln, lo, lm, L
- `M`: mn, mo, mp, M
- `N`: np, no, nq, N
- `O`: op, oq, or, O
- `P`: pq, pr, ps, P
- `Q`: qr, qs, qt, Q
- `R`: ru, rv, rw, R
- `S`: sx, sy, sz, S
- `T`: tx, ty, tz, T
- `U`: ux, uy, uz, U
- `V`: vx, vy, vz, V
- `W`: wx, wy, wz, W
- `X`: xa, xb, xc, X
- `Y`: yd, ye, yf, Y
- `Z`: za, zb, ze, Z
### APPENDIX V

Table A.3  Conversion of Raw Score to Ability Estimate

<table>
<thead>
<tr>
<th>Raw Score</th>
<th>Auditory Discrimination</th>
<th>Recognition of Initial Consonant Sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items 1–40</td>
<td>Items 1–17</td>
</tr>
<tr>
<td></td>
<td>Mean difficulty 50.0</td>
<td>Mean difficulty 50.0</td>
</tr>
<tr>
<td></td>
<td>Standard deviation 5.8</td>
<td>Standard deviation 2.8</td>
</tr>
<tr>
<td>Raw score</td>
<td>Ability estimate</td>
<td>Standard error</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>0</td>
<td>&lt;29</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>29</td>
<td>5.8</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>4.1</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>3.4</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>3.0</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>2.7</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>41</td>
<td>2.4</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>43</td>
<td>2.2</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
<td>2.1</td>
</tr>
<tr>
<td>11</td>
<td>45</td>
<td>2.0</td>
</tr>
<tr>
<td>12</td>
<td>45</td>
<td>2.0</td>
</tr>
<tr>
<td>13</td>
<td>46</td>
<td>1.9</td>
</tr>
<tr>
<td>14</td>
<td>46</td>
<td>1.9</td>
</tr>
<tr>
<td>15</td>
<td>47</td>
<td>1.9</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>1.8</td>
</tr>
<tr>
<td>17</td>
<td>48</td>
<td>1.8</td>
</tr>
</tbody>
</table>

continued p. 143
### TABLE A.3

#### Auditory Discrimination continued

**Items 1—40**

Mean difficulty 50.0

Standard deviation 5.8

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>55</td>
<td>2.0</td>
</tr>
<tr>
<td>30</td>
<td>56</td>
<td>2.1</td>
</tr>
<tr>
<td>31</td>
<td>57</td>
<td>2.2</td>
</tr>
<tr>
<td>32</td>
<td>58</td>
<td>2.3</td>
</tr>
<tr>
<td>33</td>
<td>59</td>
<td>2.4</td>
</tr>
<tr>
<td>34</td>
<td>60</td>
<td>2.5</td>
</tr>
<tr>
<td>35</td>
<td>61</td>
<td>2.7</td>
</tr>
<tr>
<td>36</td>
<td>63</td>
<td>3.0</td>
</tr>
<tr>
<td>37</td>
<td>64</td>
<td>3.4</td>
</tr>
<tr>
<td>38</td>
<td>67</td>
<td>4.1</td>
</tr>
<tr>
<td>39</td>
<td>71</td>
<td>5.8</td>
</tr>
<tr>
<td>40</td>
<td>&gt; 71</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Number Form Items 1—8

Mean difficulty 49.8

Standard deviation 6.3

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt; 38</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>6.5</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>4.3</td>
</tr>
<tr>
<td>5</td>
<td>53</td>
<td>4.5</td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>5.0</td>
</tr>
<tr>
<td>7</td>
<td>62</td>
<td>6.5</td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Figure Formation Items 1—4

Mean difficulty 50.0

Standard deviation 2.1

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt; 45</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>45</td>
<td>5.4</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>4.7</td>
</tr>
<tr>
<td>3</td>
<td>55</td>
<td>5.4</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>-</td>
</tr>
</tbody>
</table>
### Table A.3 continued

#### Prepositions
**Items 1–9**
Mean difficulty 50.0
Standard deviation 5.2

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;39</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>&lt;39</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>5.8</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>47</td>
<td>4.0</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>3.9</td>
</tr>
<tr>
<td>6</td>
<td>53</td>
<td>4.0</td>
</tr>
<tr>
<td>7</td>
<td>56</td>
<td>4.5</td>
</tr>
<tr>
<td>8</td>
<td>61</td>
<td>5.8</td>
</tr>
<tr>
<td>9</td>
<td>&gt;61</td>
<td></td>
</tr>
</tbody>
</table>

#### Verb Tense
**Items 1–9**
Mean difficulty 50.0
Standard deviation 6.7

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;37</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>37</td>
<td>6.4</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>4.8</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>4.3</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>4.0</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>4.0</td>
</tr>
<tr>
<td>6</td>
<td>54</td>
<td>4.3</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>4.8</td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>6.4</td>
</tr>
<tr>
<td>9</td>
<td>&gt;62</td>
<td></td>
</tr>
</tbody>
</table>

#### Pronouns
**Items 1–11**
Mean difficulty 50.0
Standard deviation 3.5

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;38</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>5.2</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>3.9</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>4.4</td>
</tr>
<tr>
<td>4</td>
<td>47</td>
<td>3.1</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>3.0</td>
</tr>
<tr>
<td>6</td>
<td>51</td>
<td>3.0</td>
</tr>
<tr>
<td>7</td>
<td>53</td>
<td>3.1</td>
</tr>
<tr>
<td>8</td>
<td>55</td>
<td>3.4</td>
</tr>
<tr>
<td>9</td>
<td>57</td>
<td>3.9</td>
</tr>
<tr>
<td>10</td>
<td>61</td>
<td>5.2</td>
</tr>
<tr>
<td>11</td>
<td>&gt;61</td>
<td></td>
</tr>
</tbody>
</table>

#### Negation
**Items 1–8**
Mean difficulty 50.0
Standard deviation 4.6

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;40</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>40</td>
<td>5.7</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
<td>4.4</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>3.9</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>3.8</td>
</tr>
<tr>
<td>5</td>
<td>53</td>
<td>3.9</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>4.4</td>
</tr>
<tr>
<td>7</td>
<td>60</td>
<td>5.7</td>
</tr>
<tr>
<td>8</td>
<td>&gt;60</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>61</td>
<td>5.2</td>
</tr>
<tr>
<td>10</td>
<td>&gt;61</td>
<td></td>
</tr>
</tbody>
</table>
TABLE A.3

Comprehension
Items 1—6
Mean difficulty 50.0
Standard deviation 4.4

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;42</td>
<td>--</td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td>5.7</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
<td>4.5</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>4.2</td>
</tr>
<tr>
<td>4</td>
<td>54</td>
<td>4.5</td>
</tr>
<tr>
<td>5</td>
<td>58</td>
<td>5.7</td>
</tr>
<tr>
<td>6</td>
<td>&gt;58</td>
<td>--</td>
</tr>
</tbody>
</table>
### Table A.3 continued

<table>
<thead>
<tr>
<th>Raw score</th>
<th>Ability estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt; 28</td>
<td>—</td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>6.1</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>4.4</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>3.6</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>3.2</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>2.9</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>2.7</td>
</tr>
<tr>
<td>7</td>
<td>41</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>2.4</td>
</tr>
<tr>
<td>9</td>
<td>44</td>
<td>2.3</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>2.2</td>
</tr>
<tr>
<td>11</td>
<td>45</td>
<td>2.2</td>
</tr>
<tr>
<td>12</td>
<td>45</td>
<td>2.1</td>
</tr>
<tr>
<td>13</td>
<td>46</td>
<td>2.1</td>
</tr>
<tr>
<td>14</td>
<td>47</td>
<td>2.0</td>
</tr>
<tr>
<td>15</td>
<td>47</td>
<td>2.0</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>2.0</td>
</tr>
<tr>
<td>17</td>
<td>49</td>
<td>2.0</td>
</tr>
<tr>
<td>18</td>
<td>49</td>
<td>2.0</td>
</tr>
<tr>
<td>19</td>
<td>50</td>
<td>2.0</td>
</tr>
<tr>
<td>20</td>
<td>51</td>
<td>2.0</td>
</tr>
<tr>
<td>21</td>
<td>51</td>
<td>2.0</td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>2.0</td>
</tr>
<tr>
<td>23</td>
<td>53</td>
<td>2.0</td>
</tr>
<tr>
<td>24</td>
<td>53</td>
<td>2.0</td>
</tr>
<tr>
<td>25</td>
<td>54</td>
<td>2.0</td>
</tr>
<tr>
<td>26</td>
<td>55</td>
<td>2.1</td>
</tr>
<tr>
<td>27</td>
<td>55</td>
<td>2.1</td>
</tr>
<tr>
<td>28</td>
<td>56</td>
<td>2.2</td>
</tr>
<tr>
<td>29</td>
<td>57</td>
<td>2.2</td>
</tr>
<tr>
<td>30</td>
<td>58</td>
<td>2.3</td>
</tr>
<tr>
<td>31</td>
<td>59</td>
<td>2.4</td>
</tr>
<tr>
<td>32</td>
<td>60</td>
<td>2.5</td>
</tr>
<tr>
<td>33</td>
<td>61</td>
<td>2.7</td>
</tr>
<tr>
<td>34</td>
<td>63</td>
<td>2.9</td>
</tr>
<tr>
<td>35</td>
<td>65</td>
<td>3.2</td>
</tr>
<tr>
<td>36</td>
<td>68</td>
<td>3.6</td>
</tr>
<tr>
<td>37</td>
<td>72</td>
<td>4.4</td>
</tr>
<tr>
<td>38</td>
<td>&gt; 72</td>
<td>—</td>
</tr>
</tbody>
</table>

Word Knowledge
---

Items 1–38
Mean difficulty: 50.0
Standard deviation: 6.8
ACER
Early School Series

Word Knowledge Test

Directions for Administration
Word Knowledge Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
- Test Booklets
- Score Keys
- *Early Identification and Intervention: A Handbook for Teachers and School Counsellors*

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher’s judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child’s abilities and needs.

This booklet contains general and detailed directions for the administration of the Word Knowledge Test.

Other tests contained in the series are:

- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Auditory Discrimination Test
- Receptive Language Skills
  - Comprehension Test
  - Negation Test
  - Prepositions Test
  - Pronouns Test
  - Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Word Knowledge Test Directions for Administration. Copyright © ACER 1981.
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with *Early Identification and Intervention: A Handbook for Teachers and School Counsellors*, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child’s name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.
6 Arrange for all children to be tested to visit the toilet just before the start of the test.

7 Lay out a test booklet, crayon and marker for each child.

8 Check the names on each test booklet to ensure that each child has the correct booklet.

**During Testing**

1. Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.

2. Minimize distraction during the test.

3. Read the directions for each item exactly as printed.

4. During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.

5. In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

**Scoring — General Rules**

1. Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.

2. Example items and preliminary exercises are not to be scored.

3. Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.

4. All scoring should be re-checked to ensure maximum accuracy.

5. Follow the specific directions given for scoring this test.

6. Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8"), and a crayon or pencil.

To administer the test follow the directions exactly as given below. The test may be given in one session or several sessions. If the test is to be administered in sections over a number of days, administer the two example items, A and B, at the beginning of each testing session and then proceed to the item following the one which was administered last. Distribute booklets, pencils, and markers, making sure that each child receives the booklet with his or her name on it.

Say:
'Open your books at page one. This is the page which has one star at the top.'

Administer example items as follows:

Example A
+ ARM +

Say:
'Today we are going to work with some pictures.
Put your marker under the first row of pictures.
Notice that there is a cross (+) at the beginning and at the end of the row.
Place your marker in such a way that you can see the two crosses and the pictures in that row. (demonstrate)
The pictures are LEG, ARM, EAR, and HAND.
Now take your pencil and put a cross (demonstrate + on the board) on ARM.
Put a cross on ARM.

Check that each child has marked the correct word. If necessary illustrate by marking the test page in your demonstration booklet.
When all children have finished, say:

Example B
o BALL o

Put your marker under the next row.
There is a circle at the beginning and at the end of this row.
Put a cross on BALL.
Make sure that all children make the correct response.
If necessary help a child to make the correct choice.

When all children have finished,
say:
Now you are going to work all by yourself.
I am not going to help you any more.
Put your marker under the next row of pictures.

Note: To assist marker alignment the first row of pictures on each page has a cross (+) or circle (o) on either side. (Even numbered pages show +, odd numbered pages show o.)

Item
1 DRINK Now put a cross on DRINK.
Put your marker under the next row of pictures.

2 FISH Put a cross on FISH.
Put your marker under the next row. (Repeat this instruction for each item.)
Now we'll do the next page — the page with two stars at the top.
Fold your book over. (demonstrate)
Put your marker under the first row so that you can see the cross at the beginning and at the end of the row. (demonstrate)

3 + HORSE + Put a cross on HORSE.

4 SHEEP Put a cross on SHEEP.
Check to see that each child has the correct page and line.

5 SLEEP Put a cross on SLEEP.

6 SCISSORS Put a cross on SCISSORS.
Turn your book over.
This page has three stars at the top.
Put your marker under the first row, so that you can see the circle on each side.

7 o CLOCK o Put a cross on CLOCK.
8 FAT - Put a cross on FAT.  
Check to see that every child has the correct page and line.

9 PAPER Put a cross on PAPER.

10 FINGER Put a cross on FINGER.
  
  Turn the page over.  
  Fold your book back. (demonstrate)  
  This page has four stars at the top.  
  Put your marker under the first row, so that you can see a cross on either side.

11 + COAT + Put a cross on COAT.

12 BIKE Put a cross on BIKE.

13 WINDOW Put a cross on WINDOW.

14 BROTHER Put a cross on BROTHER.
  
  Turn your book over.  
  This page has five stars at the top.  
  Put your marker under the first row, (o ... o)

15 o HAMMER o Put a cross on HAMMER.

16 MAN Put a cross on MAN.

17 PIG Put a cross on PIG.

18 STORY Put a cross on STORY.
  
  Turn the page over.  
  Fold back your book.  
  This page has six stars at the top.  
  Put your marker under the first row, (+ ... +)

19 + FIRE + Put a cross on FIRE.

20 GRASS Put a cross on GRASS.

21 BIRTHDAY Put a cross on BIRTHDAY.
22 PAINT  Put a cross on PAINT.

Turn your book over.
This page has seven stars at the top.
Put your marker under the first row. (o ... o)

23 o FLY o  Put a cross on FLY.

24 DOCTOR  Put a cross on DOCTOR.

25 SLIDE  Put a cross on SLIDE.

26 PUSH  Put a cross on PUSH.

Turn the page.
Fold back your book.
This page has eight stars at the top.
Put your marker under the first row. (+ ... +)

27 + BUCKET +  Put a cross on BUCKET.

28 LAUGH  Put a cross on LAUGH.

29 BEACH  Put a cross on BEACH.

30 BUTTON  Put a cross on BUTTON.

Turn your book over.
This page has nine stars at the top.
Place your marker under the first row. (o ... o)

31 o KNIFE o  Put a cross on KNIFE.

32 OLD  Put a cross on OLD.

33 HAPPY  Put a cross on HAPPY.

34 BRUSH  Put a cross on BRUSH.

Turn the page.
This page has ten stars at the top.
Put your marker under the first row of pictures. (+ ... +)

35 + GARDEN +  Put a cross on GARDEN.

36 FELL  Put a cross on FELL.
37 NEST  Put a cross on NEST.
38 GLASS  Put a cross on GLASS.
          Now put down your crayons and turn your books over.

Collect the booklets as soon as the final item has been completed.

Scoring (Maximum score: 38)
If the test is administered in a number of sessions, leave the scoring until
the whole test has been completed.
Do not score example items.
- Score one point for each correct item crossed, i.e. possible scores for
each item are zero (0) or one (1). Zero (0) is given for items in which
either one or more wrong pictures, all pictures, or none of the pictures
have been crossed.
During scoring, place a tick (v) in the right-hand margin next to each
correct item. Count up the number of correct items and record the total
number of correct items where indicated on the front of the test booklet.
ACER
Early
School
Series

Comprehension Test

Directions for Administration

Australian Council for Educational Research
Frederick Street, Hawthorn 3122, Victoria

Copyright © ACER 1981
Comprehension Test
Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
Test Booklets
Score Keys
Early Identification and Intervention:
A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Comprehension Test.

Other tests contained in the series are:

- Auditory Discrimination Test
- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
  - Negation Test
  - Prepositions Test
  - Pronouns Test
  - Verb Tense Test

*Early Identification and Intervention: A Handbook for Teachers and School Counsellors* is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Comprehension Test. Directions for Administration. Copyright © ACER 1981
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child's name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.

ACER Early School Series. Comprehension Test Directions for Administration Copyright © ACER 1981
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.

ACER Early School Series. Comprehension Test Directions for Administration. Copyright © ACER 1981
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5” x 8”), and a crayon or pencil.

To administer the test follow the directions as given below. It is essential that each item is read exactly as printed.

Distribute the test booklets, pencils and markers, making sure that each child receives the booklet with his or her name on it.

Wait until all children are ready to begin.

Say:

Now (Today) I am going to tell you some stories.
You will find the picture that shows what I tell you.
Open your book to the page that has one star at the top. (demonstrate)
Put your marker under the top row of pictures like this. (demonstrate)

Check that all children have positioned their markers correctly.

Say:
Now listen to this story.

Practice item A TV

The children like watching TV. Sometimes the picture is not very good. There is something wrong with the TV. Father has to fix it. He has fixed it. Now all the family can watch TV again.
Put a cross on the picture that shows the family watching TV.

Check that all children have placed a cross on the correct picture, and assist those who are finding the task difficult.

Make sure that all children have understood what is required. If necessary point to each picture and explain.
Here the boy is watching TV. In this picture the girl is sitting in front of TV. I think something has gone wrong... In this picture father checks whether he has fixed the picture... Now point to the picture which shows all the family watching TV again... Yes,
that's right. Now put a cross (demonstrate x on the board) on the picture which shows the family watching TV.

Give no further help.
Say:
Now put your marker under the next row of pictures.
The middle row.

Check that all children have positioned their markers correctly.
Say:

Item 1
POOL
Jim and Anne like to go to the swimming pool. They always take their bathers and towels as they love to swim. They splash in the water and jump off the edge. Put a cross on the picture that shows this.

Put your marker under the bottom row of pictures.

Check that all children have moved their markers to the required position.

Say:

Item 2
WORM
Baby bird was hungry. He sat on a branch near his nest. Mother bird brought him a fine worm and put it in his mouth.

Wait until all children have finished.

Say:

Turn the page over.
This page has two stars at the top.
Put your marker under the top row.

Item 3
PIZZA
Make a cross on this picture. It is made from dough. Cheese, tomatoes and salami make it good to eat. It is cooked in a big oven in a special shop. People cut it in pieces and eat it.

Put your marker under the next row of pictures.
Item 4

DOLL

Put a cross on the picture that shows this:
The little doll is on the bed.
Two girls are standing by the bed.
Put a cross on that picture.
Put your marker under the next row.

Check that all children have positioned their markers correctly.
Say:

Item 5

HOLIDAY

Make a cross on this picture.
Bill and Joan are going for a holiday.
They are to travel by train.
Mother and father are on the station with them.
They have a case for their clothes and a bag for their toys.

Put your marker under the bottom row of pictures.

Item 6

HOUSE

Make a cross on this picture.
John and Mary were playing with blocks.
They built a large house.
It had a chimney, two windows and a door.

Put your pencil down and turn your book over.
You can have a little rest now.

Collect the test booklets as soon as the last item has been completed.

Scoring (Maximum score: 6)
Score one point for each correct item. Possible scores for each item are one (1) and zero (0). Zero (0) is given for items in which an incorrect picture, no picture, or more than one picture has been crossed. For an item in which one correct and one or more incorrect pictures were marked, zero (0) points are given.

During scoring, place a tick (✓) for each correct item in the appropriate boxes provided for this purpose in the right-hand margin.

Record the total number of correct responses where indicated on the front of the test booklet.

ACER Early School Series. Comprehension Test. Directions for Administration. Copyright © ACER 1981
Negation Test

Directions for Administration
ACER Early School Series

Negation Test
Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
  Test Booklets
  Score Keys
  Early Identification and Intervention
  A Handbook for Teachers and School Counsellors

Copyright © ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher’s judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child’s abilities and needs.

This booklet contains general and detailed directions for the administration of the Negation Test.

Other tests contained in the series are:

- Auditory Discrimination Test
- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
- Comprehension Test
- Prepositions Test
- Pronouns Test
- Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Negation Test Directions for Administration Copyright © ACER 1981
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.

2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.

3. Prior to the day of testing, enter the child’s name and other details required on the front of each test booklet.

4. Have available a sufficient supply of crayons or pencils and markers for each child.

5. Have available a box of tissues.

6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.

7. Attempt to obtain the help of another teacher as assistant for the days of testing.

8. Arrange place for testing.

9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing

1. Have this booklet open at the pages giving the detailed directions.

2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.

   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.

3. Have available paper and pencil for yourself in case it is necessary for you to make notes.

4. See that the tables or desks during the testing are cleared.

5. Arrange seating in such a way as to avoid copying.

ACER Early School Series, Nogation Test Directions for Administration. Copyright © ACER 1981
Arrange for all children to be tested to visit the toilet just before the start of the test.

Lay out a test booklet, crayon and marker for each child.

Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing

1. Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2. Minimize distraction during the test.
3. Read the directions for each item exactly as printed.
4. During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5. In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules

1. Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2. Example items and preliminary exercises are not to be scored.
3. Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after 'Number of correct items' on the front of each test booklet.
4. All scoring should be re-checked to ensure maximum accuracy.
5. Follow the specific directions given for scoring this test.
6. Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8") and a crayon or pencil.

To administer the test follow the directions as given below. It is essential that each item be read exactly as printed. Distribute the test booklets, pencils and markers, making sure that each child receives the correct booklet, i.e. the booklet with his or her name on it.

Wait until the children are ready to begin.
Say:
Today we are going to work with more pictures.
I will tell you something about one of these pictures, and I want you to put a cross (demonstrate on the board) on the picture I am talking about.

Do not open your booklet.
Put your marker under the row of pictures at the bottom of the page.

Check that all children have followed this direction.
Say:
Practice item A
They have built a house without a door.
Put a cross on that picture.

Check that all children have placed a cross on the correct picture, and help those children who have found the task difficult.

Wait until all children have understood the task.
Say:
Turn to the page with one star at the top.
Put your marker under the first row. (demonstrate)

Note: Do not give any further help.
Let's look at these pictures.
Find the one I tell you about.
The cake has no candles on it.
Make a cross on that picture.

Put your marker under the next row.
The hen hasn't any chicks.
Make a cross on that picture.

Put your marker under the next row.
The bear couldn't get through the fence.
Make a cross on that picture.

Put your marker under the next row.
The fish is not in the water.
Make a cross on that picture.

Now turn the page, like this. (demonstrate)
This page has two stars at the top.

Check that each child has the correct page.
Say:

Put your marker under the top row.
The girl is not happy.
Make a cross on that picture.

Put your marker under the next row.
The car won't go.
Make a cross on that picture.

Put your marker under the next row.
Find the one that isn't a boy.
Make a cross on that picture.

Put your marker under the next row.
The boy can't reach the shelf.
Make a cross on that picture.

Put your pencils down. Turn your book over.
Have a rest now.

Collect the test booklets as soon as the last item has been completed.
Scoring (Maximum score: 8)
Score one point for each correct item. Possible scores for each item are one (1) and zero (0). Zero (0) is given for items in which an incorrect picture, no picture, or more than one picture has been crossed. For an item in which one correct and one or more incorrect pictures were marked, zero (0) points are given.

During scoring place a tick (✓) for each correct item in the appropriate boxes provided for this purpose in the right-hand margin.

Record the total number of correct responses where indicated on the front of the test booklet.
Verb Tense Test

Directions for Administration
Verb Tense Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
  Test Booklets
  Score Keys
  Early Identification and Intervention: A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Verb Tense Test.

Other tests contained in the series are:
- Auditory Discrimination Test
- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
- Comprehension Test
- Negation Test
- Prepositions Test
- Pronouns Test

*Early Identification and Intervention: A Handbook for Teachers and School Counsellors* is common to all 10 tests, and contains information concerning, the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series, Verb Test Directions for Administration. Copyright © ACER 1981.
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with *Early Identification and Intervention: A Handbook for Teachers and School Counsellors*, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child’s name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible, no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleaned.
5. Arrange seating in such a way as to avoid copying.

ACER Early School Series *Vita Test Directions for Administration* Copyright © ACER 1981
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8") and a crayon or pencil.

To administer the test follow the directions as given below. It is essential that each item is read exactly as printed. Distribute the test booklets, pencils, and markers, making sure that each child receives the booklet with the correct name on it.

Wait until the children are ready to begin.

Say:
Open your book so that this page, the page with one star at the top, is in front of you. Fold back the page. (demonstrate folding the booklet)
This is a picture story about a bear family.

Point to the picture at the top of the page.
Say:
Here are three bears, mother bear, father bear, and baby bear.
Let's see what they are doing.
I will tell you what they are doing, and I want you to put a cross (demonstrate) on the picture which shows what I tell you about the bears.
Ready?

Put your marker under the first row of bear pictures, like this. (demonstrate)

Practice item A

Baby bear is running to the house.
Put a cross on that picture.

Check that all children have put a cross on the correct picture.
If a child shows confusion, give assistance.

ACER Early School Series. Verb Test. Directions for Administration. Copyright © ACER 1981
Say:
Look! Here baby bear is standing outside the house; here he is sitting on the swing; here he is walking through the gate; and what is he doing here? (wait for a reply)
Yes, he is running to the house.
You put a cross on that picture.

Wait until all children have understood the task.

Note: Give no further help. Repeat each item once, only if required.

Say:
Now you will work by yourself.
I am not going to help you any more.
Listen carefully and put a cross on the picture I am telling you about.

Put your marker under the next row of pictures.

Item 1 Baby bear has gone to bed.
Put a cross on that picture.

Item 2 Put your marker under the next row.
Baby bear fell down the stairs.
Put a cross on that picture.

Turn over your book. (demonstrate)
This page has two stars at the top.

Check that each child has the correct page.
Say:
Item 3 Put your marker under the top row.
Baby bear has eaten his cake.
Put a cross on that picture.

Item 4 Put your marker under the next row.
Baby bear is chasing the cat.
Put a cross on that picture.

ACER Early School Series. Verb Test. Directions for Administration. Copyright © ACER 1981
Item 5  
Put your marker under the next row.
The cat is chasing baby bear.
Put a cross on that picture.

Item 6  
Put your marker under the next row.
All the biscuits have been eaten.
Put a cross on that picture.

Turn the page.
Fold over the book. (demonstrate)
This page has three stars at the top.

Check that each child faces the correct page.
Say:

Item 7  
Put your marker under the top row.
Baby bear has finished his milk.
Put a cross on that picture.

Item 8  
Put your marker under the next row.
Father bear has closed the gate.
Put a cross on that picture.

Item 9  
Put your marker under the next row.
Father bear will answer the telephone.
Put a cross on that picture.

Put down your pencil.
Turn over your book.
You can have a little rest now.

Collect test booklets as soon as the last item has been completed.

Scoring (Maximum score: 9)
Score one point for each correct item. Possible scores for each item are one (1) and zero (0). Zero (0) is given for items in which an incorrect picture, no picture, or more than one picture has been crossed. For an item in which one correct and one or more incorrect pictures were marked, zero (0) points are given.

During scoring, place a tick (✓) for each correct item in the appropriate boxes provided for this purpose in the right-hand margin.

Record the total number of correct responses where indicated on the front of the test booklet.

ACER Early School Series. Verb Test Directions for Administration. Copyright © ACER 1981
Pronouns Test

Directions for Administration
Pronouns Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
- Test Booklets
- Score Keys
- Early Identification and Intervention;
- A Handbook for Teachers and School Counsellors

Copyright © ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Pronouns Test.

Other tests contained in the series are:
- Auditory Discrimination Test
- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
  - Comprehension Test
  - Negation Test
  - Prepositions Test
  - Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child’s name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at one time.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.

ACER Early School Series. Pronouns Test Directions for Administration. Copyright © ACER 1981
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8") and a crayon or pencil.

To administer the test follow the directions as given below. It is essential that each item is read exactly as printed. Distribute the test booklets, pencils, and markers, making sure that each child receives the booklet with his or her name on it.

Wait until the children are ready to begin.
Say:
Now (Today) we are going to look at some pictures about the family.
Here is a mother, a father, a boy and a girl.
(Point to the picture of the family group)
Let's see what they are doing.
Put your marker under the row of pictures at the bottom of the page.

Demonstrate and help those children who may find the task difficult.
Say:
Practice item A
The family is looking at books.
Father has a book, the boy has a book, the girl has a book.
Put a cross (demonstrate x on the board) on hers.

Check that all children have placed a cross on the correct picture, and assist those who are finding the task difficult.

Make sure that all children have understood what is required. If necessary, point to each picture and explain:

This is his book, this is her book, this is their book, this is his. Now point to hers.... That's right... Now put a cross on hers.

Give no further help.

ACER Early School Series. Pronouns Test Directions for Administration. Copyright © ACER 1961
Say:
Open your book.
Fold back the page. (demonstrate)
This page has one star at the top.

Demonstrate and make sure that each child has found the required page.
Say:
Put your marker under the first row of pictures. (demonstrate)
Listen carefully to what I say.

Item 1
The dog chased her.
Make a cross on that picture.

Put your marker under the next row of pictures.

Item 2
Her hands are dirty.
Put a cross on that picture.

Put your marker under the next row.

Check that each child’s marker is in the correct position for item 3.
Say:

Item 3
They are looking at TV.
Put a cross on that picture.

Put your marker under the bottom row of pictures.

Item 4
Mother gives a cake to him.
Put a cross on that picture.

Now turn your book over. This page has two stars at the top. (demonstrate)

Put your marker under the top row of pictures.

Check that each child has found the required page.
Item 5  She has roller skates.
Make a cross on that picture.
Put your marker under the next row.

Item 6  The boy sees them.
Put a cross on that picture.
Put your marker under the next row.
Check that all children have placed their markers under item 7.
Say:

Item 7  He is walking.
Put a cross on that picture.
Put your marker under the bottom row of pictures.

Item 8  It is their birthday.
Put a cross on that picture
Turn the page. Fold over the book.
This page has three stars at the top.
Put your marker under the first row of pictures.
Check that all children have positioned their markers correctly.
Say:

Item 9  We like icecream.
Put a cross on that picture.
Put your marker under the next row.

Item 10  It's wheel has gone.
Make a cross on that picture.
Put your marker under the last row of pictures.

Item 11  His balloon is big.
Put a cross on that picture.

Put your pencil down and turn your book over. Have a rest now.

ACER Early School Series. Pronouns Test Directions for Administration. Copyright © ACER 1981
Collect the test booklets as soon as the last item has been completed.

Scoring (Maximum score: 11)
Score one point for each correct item. Possible scores for each item are one (1) and zero (0). Zero (0) is given for items in which an incorrect picture, no picture, or more than one picture has been crossed. For an item in which one correct and one or more incorrect pictures were marked, zero (0) points are given.

During scoring, place a tick (✓) for each correct item in the appropriate boxes provided for this purpose in the right-hand margin.

Record the total number of correct responses where indicated on the front of the test booklet.
Prepositions Test

Directions for Administration
Prepositions Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
Test Booklets
Score Keys
Early Identification and Intervention:
A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Prepositions Test.

Other tests contained in the series are:
- Auditory Discrimination Test
- Figure Formation Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
- Comprehension Test
- Negation Test
- Pronouns Test
- Verb Tense Test

*Early Identification and Intervention: A Handbook for Teachers and School Counsellors* is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Prepositions Test Directions for Administration. Copyright © ACER 1981
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with *Early Identification and Intervention: A Handbook for Teachers and School Counsellors*, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.

2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.

3. Prior to the day of testing, enter the child's name and other details required on the front of each test booklet.

4. Have available a sufficient supply of crayons or pencils and markers for each child.

5. Have available a box of tissues.

6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.

7. Attempt to obtain the help of another teacher as assistant for the days of testing.

8. Arrange place for testing.

9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing

1. Have this booklet open at the pages giving the detailed directions.

2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.

   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet so that the children can see exactly what you are referring to.

3. Have available paper and pencil for yourself in case it is necessary for you to make notes.

4. See that the tables or desks used during the testing are cleared.

5. Arrange seating in such a way as to avoid copying.

ACER Early School Series. Prepositions Test Directions for Administration. Copyright © ACER 1981
6. Arrange for all children to be tested to visit the toilet just before the start of the test.
7. Lay out a test booklet, crayon and marker for each child.
8. Check the names on each test booklet to ensure that each child has the correct booklet.

**During Testing**
1. Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2. Minimize distraction during the test.
3. Read the directions for each item exactly as printed.
4. During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5. In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

**Scoring — General Rules**
1. Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2. Example items and preliminary exercises are not to be scored.
3. Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4. All scoring should be re-checked to ensure maximum accuracy.
5. Follow the specific directions given for scoring this test.
6. Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8") and a crayon or pencil.

To administer the test follow the directions as given below. It is essential that each item be read exactly as printed. Distribute the test booklets, pencils, and markers, making sure that each child receives the booklet with his or her name on it.

Wait until the children are ready to begin.
Say:
Open your book.
Fold back the page. (demonstrate)
This page has one star at the top.

Demonstrate and help those children who are finding the task difficult.

Point to the cat at the top of the page.
Say:
Here is Sooty the cat.
He is having fun.
He does a lot of things.
We are going to play a game with Sooty.

Put your marker under the first row of pictures, like this. (demonstrate)
I will tell you what Sooty does, and
I want you to put a cross (x) like this
(demonstrate on the board or sample booklet)
on the picture which shows what I say.
Now listen carefully.

Practice item A
The cat is on the chair.
Put a cross on that picture.

Check that all children have placed a cross on the correct picture, and assist those children who are finding the task difficult.
Wait until all children have understood the task.

Say:
Yes, Sooty is on the chair in this picture.
(demonstrate by holding up the test booklet and by marking the correct picture)

Proceed to Practice item B.
Say:
Now put your marker under the next row.
The cat is jumping over the fence.
Put a cross on that picture.

Again check that all children have marked the correct picture. Give help where needed.

Note: From this point on, no further help should be given with the test items. Repeat each item only once unless there is a gross distraction which might prevent the child from hearing accurately, e.g. a person entering the room, public address announcements, etc.

It is essential that each item is read exactly as printed.

Now you will work by yourself.
I am not going to help you any more.
Listen carefully and put a cross on the right picture.
Ready?

Wait until all children are ready to proceed.
Say:
Put your marker under the next row.
The cat is up the tree
Put a cross on that picture.

Now turn the book over.
This is page two.
It has two stars at the top.
(point to the stars)

Check that each book is open at the page with two stars at the top.
Say:

Item 2
Put your marker under the top row.
The cat is jumping through the window.
Put a cross on that picture.

Item 3
Put your marker under the next row.
The cat is next to the fire.
Put a cross on that picture.

Item 4
Put your marker under the next row.
The cat is outside the house.
Put a cross on that picture.

Item 5
Put your marker under the next row.
The cat is behind the tree.
Put a cross on that picture.

Turn the page.
Fold your book. (demonstrate)
This page has three stars at the top.

Make sure that each child has the correct page.
Say:
Put your marker under the top row.

Item 6
The cat is between the boxes.
Put a cross on that picture.

Item 7
Put your marker under the next row.
The cat is beside the blocks.
Put a cross on that picture.

Item 8
Put your marker under the next row.
The cat is with the dog.
Put a cross on that picture.

Item 9
Put your marker under the next row.
The cat runs after the mouse.
Put a cross on that picture.
That's all about Sooty.
Put your pencils down and close your books.
Have a little rest now.

Collect the test booklets as soon as the last item has been completed.
Scoring (Maximum score: 9)
Score one point for each correct item. Possible scores for each item are one (1) and zero (0). Zero (0) is given for items in which an incorrect picture, no picture, or more than one picture has been crossed. For an item in which one correct and one or more incorrect pictures were marked, zero (0) points are given.

During scoring, place a tick (✓) for each correct item in the appropriate boxes provided for this purpose in the right-hand margin.

Record the total number of correct responses where indicated on the front of the test booklet.
Figure Formation Test

Directions for Administration
Figure Formation Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
- Test Booklets
- Score Keys
- Early Identification and Intervention:
  A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners' grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Figure Formation Test.

Other tests contained in the series are:
- Auditory Discrimination Test
- Number Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
  - Comprehension Test
- Negation Test
- Prepositions Test
- Pronouns Test
- Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Figure Formation Test Directions for Administration. Copyright © ACER 1981
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child’s name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.

ACER Early School Series. Figure Formation Test Directions for Administration. Copyright © ACER 1981.
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after 'Number of correct items' on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet and a crayon or pencil.

To administer the test follow the directions exactly as given here. After each item, wait until all children have finished.

Place a test booklet in front of each child so that the front page is face up. Make sure that each child has the necessary test materials. Wait until all children are ready to begin, then hold up the test booklet and say to the children:

Open your books at page one.
This is the page which has one star at the top.

Point to the star at the top of the page. Wait until all children have turned to the correct page.
Say:

Example

BOAT
On this page we have a boat and underneath it many parts which make up the boat.
Look at the boat carefully, then look at the parts.
Now put a cross on all the parts which are needed to make up the boat.

Wait until all children have finished. Then check that each child has crossed out the correct parts. Assist children who have been unable to cross out the correct parts.
Say:

Turn the page to page two.
This is the page which has two stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:

Item 1
TRUCK
On this page we have a truck and underneath it many parts which make up the truck.
Look at the truck carefully, then look at the parts.
Now put a cross on all the parts which are needed to make up the truck.

Wait until all children have finished.
Say:
Turn the page over to the next picture.
This is the page which has three stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:
On this page we have a teapot and underneath it many parts which make up the teapot.
Look at the teapot carefully, then look at the parts.
Now put a cross on all the parts which are needed to make up the teapot.
Wait until all children have finished.
Say:
Turn to the next page.
This is the page which has four stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:
On this page we have a robot and underneath it many parts which make up the robot.
Look at the robot carefully, then look at the parts.
Now put a cross on all the parts which are needed to make up the robot.
Wait until all children have finished.
Say:
Turn the page over to the next picture.
This is the page which has five stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:
On this page we have a church and underneath it many parts which make up the church.
Look at the church carefully, then look at the parts.
Now put a cross on all the parts which are needed to make up the church.
Wait until all children have finished.
Say:
Now put down your crayons and turn over your books very quickly.
Collect the test booklets as soon as the last item has been completed.

Scoring (Maximum score: 4)
Score one point for each correctly completed item, i.e. a child can receive a score of zero (0) or one (1) for each item.
Place a tick (✓) near the bottom corner of each correct item. The child’s total score for this test is the sum of completely correct items.
Record the total score where indicated on the front of the test booklet.

<table>
<thead>
<tr>
<th>Item</th>
<th>Number of correct parts</th>
<th>Number of incorrect parts</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex.</td>
<td>BOAT</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>TRUCK</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>TEAPOT</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>ROBOT</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>CHURCH</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
Number Test

Directions for Administration
Number Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
- Test Booklets
- Score Keys
- Early Identification and Intervention: A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher’s judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child’s abilities and needs.

This booklet contains general and detailed directions for the administration of the Number Test.

Other tests contained in the series are:
- Figure Formation Test
- Auditory Discrimination Test
- Recognition of Initial Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
- Comprehension Test
- Negation Test
- Prepositions Test
- Pronouns Test
- Verb Tense Test

_Early Identification and Intervention: A Handbook for Teachers and School Counsellors_ is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning

To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child's name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing

1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.

A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored, in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
   All scoring should be re-checked to ensure maximum accuracy.
4 Follow the specific directions given for scoring this test.
5 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5” x 8” ) and a crayon or pencil.

To administer the test follow the directions exactly as given here. After each item, wait until all children have finished before proceeding to the next item.

To make sure that each child is able to produce a ‘stick’ (vertical line) of the appropriate size, three boxes for practice purposes have been provided on the back page of the test booklet.

Place the test booklet in front of each child so that the practice boxes are face up. The largest box should be on the left. Make sure that each child has the necessary test materials.

Say:

A BALLS
Today I want to see how well you can count.
Put down your crayon and count these little balls.

Draw three balls (●●●) on the blackboard.

Say:
How many balls did I draw? (pause for response) Yes, there are three balls.

Point to each ball in turn.

Say:
One, two, three — there are three balls on the blackboard. So we make three sticks in this box.

Hold up test booklet and demonstrate.

Now pick up your crayons and make three sticks in your first box — one stick for each ball.

Make sure that each child has placed three sticks in the largest box. Assist those children who have not recorded the three sticks correctly.

Say:
Now, put down your crayons.

ACER Early School Series. Number Test Directions for Administration. Copyright © ACER 1981.
While drawing two circles (○○) on the blackboard, say:
Now count these circles. How many circles did I draw?
Pause for response. Say:
Yes, two circles.
Hold up the test booklet and while drawing two sticks in the middle box, say:
There are two circles on the blackboard.
So you make two sticks in the box.
Now pick up your crayon and in the middle box make as many sticks as there are circles.
One stick for each circle.
Allow children time to finish this task, then check each child’s work and assist those children who have not recorded the two sticks correctly.
Say:
Now put down your crayons and watch what I do.

While drawing seven crosses (× × × × × × ×) on the blackboard:
Say:
Count these crosses and in the last box make as many sticks as there are crosses.
One stick for each cross. Pick up your crayon and do it now.
Pause until most children have finished this task, then, pointing to each cross in turn, say:
One, two, three, four, five, six, seven — there are seven crosses on the blackboard. So we make seven sticks in this box.
Hold up the test booklet and demonstrate. Check each child’s work and assist those children who have not recorded the seven sticks correctly.

ACER Early School Series. Number Test Directions for Administration. Copyright © ACER 1981.
Say:
Now, put down your crayons and turn your books over. (pause) Open your books to page one. Fold back the empty page. (demonstrate)
There is one star at the top of this page.

Hold up test booklet and point to the star at the top of the page.
Say:
From now on I want you to work very quietly, all by yourself:

Say:
Example
Put your marker under the row of dogs like this. (demonstrate)
Count the dogs silently in your head.
Don't tell anyone how many there are.
Keep it a secret. Put your finger on the box next to the dogs.

Hold up test booklet and point to the correct box.
Say:
Pick up your crayon, and in this box make as many sticks as there are dogs.
One stick for each dog.

Check each child’s work and assist those children who have not recorded the four sticks correctly.
Say:
Turn the page to page 2. This is the page which has two stars at the top. Fold back your book.

Point to the stars at the top of the page. Wait until all children have turned to the correct page and folded their books.

Say:
Item 1
Now put your marker under the row of PLANES.
(demonstrate)
Put your finger on the box next to the PLANES.
(demonstrate)
In this box make one stick for each PLANE.

Wait until all children have finished.

ACER Early School Series. Numer Test Directions for Administration. Copyright © ACER 1981.
Item 2
HOUSES

Say:
Now put your marker under the row of HOUSES.
(demonstrate)
Put your finger on the box next to the HOUSES.
(demonstrate)
In this box make one stick for each HOUSE.

Wait until all children have finished.
Say:
Turn the book over to page 3.
This is the page which has three stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:

Item 3
PILES OF MONEY

Now put your marker under the row of PILES OF MONEY. (demonstrate)
Put your finger on the box next to the PILES OF MONEY.
In this box make one stick for each PILE OF MONEY.

Wait until all children have finished.
Say:

Item 4
TRUCKS

Now put your marker under the row of TRUCKS. (demonstrate)
Put your finger on the box next to the TRUCKS. (demonstrate)
In this box make one stick for each TRUCK.

Wait until all children have finished.
Say:
Turn the page over to page 4.
This is the page which has four stars at the top. Fold over your book.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:

Item 5
CARS

Now put your marker under the row of CARS. (demonstrate)
Put your finger on the box next to the CARS. (demonstrate)
In this box make one stick for each CAR.
Wait until all children have finished.
Say:
Now put your marker under the row of MEN.
(demonstrate)
Put your finger on the box next to the MEN.
(demonstrate)
In this box make one stick for each MAN.

Wait until all children have finished.
Say:
Turn over your book. This is page 5.
This page has five stars at the top.
Point to the stars at the top of the page. Wait until
all children have turned to the correct page.
Say:

Now put your marker under the row of
PAIRS OF SHOES. (demonstrate)
Put your finger on the box next to the
PAIRS OF SHOES. (demonstrate)
In this box make one stick for each
PAIR OF SHOES.

Wait until all children have finished.
Say:

Now put your marker under the row of
GROUPS OF LADIES. (demonstrate)
Put your finger on the box next to the
GROUPS OF LADIES. (demonstrate)
In this box make one stick for each
GROUP OF LADIES.

Wait until all children have finished.
Say:
Put your crayons down and close your
books very quickly.

Collect the booklets as soon as the final item has been completed.
**Scoring (Maximum score: 8)**

Score one point for each correctly-completed item, i.e., a child can receive a score of zero (0) or one (1) for each item.

Place a (✓) near the answer box for each correct item. The child’s total score for this test is the sum of correct items.

Record the total score where indicated on the front of the test booklet.

---

**Answer Key**

<table>
<thead>
<tr>
<th>Item</th>
<th>Correct number of sticks</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex. dogs</td>
<td>4</td>
<td>not scored</td>
</tr>
<tr>
<td>1 planes</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>2 houses</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>3 piles of money</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>4 trucks</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>5 cars</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>6 men</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>7 pairs of shoes</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>8 groups of ladies</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Count 'sticks' which have been placed outside the answer box as part of the child’s response only if they are in the vicinity of the box provided for the item.

The child who has completed an item by placing one stick above or below each house, plane, etc., does not receive credit for this item unless the answer box also contains the correct number of sticks.
Recognition of Initial Consonant Sounds Test

Directions for Administration
Recognition of Initial Consonant Sounds Test

Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
  Test Booklets
  Score Keys
  Early Identification and Intervention:
  A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher’s judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child’s abilities and needs.

This booklet contains general and detailed directions for the administration of the Recognition of Initial Consonant Sounds Test.

Other tests contained in the series are:
- Figure Formation Test
- Number Test
- Auditory Discrimination Test
- Word Knowledge
- Receptive Language Skills
  - Comprehension Test
  - Negation Test
  - Prepositions Test
  - Pronouns Test
  - Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981.
GENERAL DIRECTIONS FOR ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with Early Identification and Intervention: A Handbook for Teachers and School Counsellors, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child's name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8"), and a crayon.

To administer the test follow the directions exactly as given here. The test may be given in one session or broken up and given in two or three sessions. If the test is to be administered in sections over a number of days, administer Preliminary Exercise 1 and the Example Item 1 at the beginning of each testing session and then proceed to the item following the one which was administered last.

Do not elaborate on the pictures of the test items. Make sure there is no interfering noise. Speak clearly and distinctly. Do not repeat any stimulus word unless you have made an error. Wait after each item until all children have finished.

Throughout this test give only the letter SOUND of the initial consonant. Do not use letter names.

Place a test booklet in front of each child so that the front page is face up. Make sure that each child has the necessary test materials. Wait until all children are ready to begin.

Preliminary Exercise 1

Say:

BALL
BOOK
BOAT
BEAR

Put your marker under your first row of pictures, like this. (demonstrate) (pause)

Put your finger on each picture as I say its name:

BALL ... BOOK ... BOAT ... BEAR.

They all start with 'b'. (sound the letter, do not use the letter-name)

Say 'b', (let children say it)

Close your lips tightly, now open them and say 'b'. (let children say it)

Listen again for the 'b' as I say the words.

BALL (pause), BOOK (pause), BOAT (pause), BEAR (pause).

Point to each picture again and say its name after me:

BALL (pause), BOOK (pause), BOAT (pause), BEAR (pause).

These words all start with 'b'.

Say 'b'.
Example

Say:

BONE
TENT
BAG
MAP

Now put your marker under the next row of pictures like this. (demonstrate)
Point to each picture.

Say:

Put your finger on each picture as I say its name:
BONE ... TENT ... BAG ... MAP.
Now point to each picture again and say its name after me:
BONE (pause), TENT (pause), BAG (pause), MAP (pause).

Do any words start with 'b'. (pause for response)
Yes, BONE and BAG start with 'b'.
Point to BONE and say BONE.
Point to BAG and say BAG.
Put a cross on BONE and a cross on BAG because they start with 'b'.

Like this. (demonstrate)

Check that each child has crossed out the correct pictures and assist those children who have not. Then proceed to Item 1.

Say:

From now on we are going to keep our crosses a secret.
Don't let anyone see where you put your crosses.
Don't peep at anyone's book.
Now put your marker under the next row of pictures, like this. (demonstrate)

Point to each picture.

Say:

Item 1

Put your finger on each picture as I say its name:
TELEPHONE
BIRD
DOG
BIKE
TELEPHONE ... BIRD ... DOG ... BIKE.

Now point to each picture again and say its name after me:
TELEPHONE (pause), BIRD (pause), DOG (pause), BIKE (pause).

Put a cross on the pictures that start with 'b'.
Now remember, it's a secret.

Wait until all children have finished.

Say:

Turn the page to page 2.
This is the page which has two stars at the top.
Point to the stars at the top of the page. Wait until all children have turned to the correct page.

Preliminary exercise 2

Put your marker under your first row of pictures, like this. (demonstrate)

FLAG
FRUIT
FLOWER
FOOT

They all start with 'F'.

(sound the letter, do not use the letter-name)

Now point to each picture and say its name after me:

FLAG (pause), FRUIT (pause), FLOWER (pause), FOOT (pause).

They all start with 'F'.

Item 2

Now put your marker under the next row of pictures, like this. (demonstrate)

FIRE
FISH
KITE
NEST

Now point to each picture and say its name after me:

FIRE (pause), FISH (pause), KITE (pause), NEST (pause).

Put a cross on the pictures that start with 'F'.

Remember, it's a secret.

Wait until all children have finished.

Say:

Preliminary exercise 3

Put your marker under the next row of pictures, like this. (demonstrate)

MAN
MAT
MUG
MARBLES

They all start with 'M'.

Now point to each picture and say its name after me:

MAN (pause), MAT (pause), MUG (pause), MARBLES (pause).

They all start with 'M'.

Say:

Item 3

Now put your marker under the next row of pictures, like this. (demonstrate)

MOON
HOUSE
TREE
MILK

Now point to each picture and say its name after me:

MOON (pause), HOUSE (pause), TREE (pause), MILK (pause).

Put a cross on the pictures that start with 'M'.
Wait until all children have finished.

Say:

Preliminary exercise 4
LAMB
LETTUCE
LIGHT
LOLLY

Put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
LAMB...LETTUCE...LIGHT...LOLLY.
They all start with 'I'.

Now point to each picture and say its name after me:
LAMB (pause), LETTUCE (pause), LIGHT (pause), LOLLY (pause).
They all start with 'I'.

Item 4
DRESS
HAND
LION
LADDER

Now put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
DRESS...HAND...LION...LADDER.

Now point to each picture and say its name after me:
DRESS (pause), HAND (pause), LION (pause), LADDER (pause).
Put a cross on the pictures that start with 'I'.

Wait until all children have finished.
Say:

Turn your book over to page 3.
This is the page which has three stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.

Preliminary exercise 5
JAIL
JEEP
JAR
JEANS

Put your marker under the first row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
JAIL...JEEP...JAR...JEANS.
They all start with 'J'.

Now point to each picture and say its name after me:
JAIL (pause), JEEP (pause), JAR (pause), JEANS (pause).
They all start with 'J'.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981.
Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

MOUSE ... JUG ... TABLE ... JELLY.

Now point to each picture and say its name after me:

MOUSE (pause), JUG (pause), TABLE (pause), JELLY (pause).

Put a cross on the pictures that start with 'j'. Remember, it's a secret.

When all children have finished, say:

Put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

WATER ... WEB ... WASHING ... WALL.

They all start with 'w'.

Now point to each picture and say its name after me:

WATER (pause), WEB (pause), WASHING (pause) WALL (pause).

They all start with 'w'.

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

WATCH ... SAW ... BASKET ... WINDOW.

They all start with 'w'.

Now point to each picture and say its name after me:

WATCH (pause), SAW (pause), BASKET (pause), WINDOW (pause).

Put a cross on the pictures that start with 'w'.

When all children have finished, say:

Put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

SHELF ... SHELL ... SHOVEL ... SHIRT.

They all start with 'sh'.

(produce sounds, do not use letter-names)

Now point to each picture and say its name after me:

SHELF (pause), SHELL (pause), SHOVEL (pause), SHIRT (pause).

They all start with 'sh'.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981
Item 7

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

SHOE
MATCH
FORK
SHIP

Now point to each picture and say its name after me:

SHOE (pause), MATCH (pause), FORK (pause),
SHIP (pause).

Put a cross on the pictures that start with 'sh'.
Remember, it's a secret.

Wait until all children have finished.
Say:

Turn the page over to page 4.
Fold your book back.
This is the page which has four stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:

[d]

Preliminary exercise 8

Put your marker under the first row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

DONKEY
DRINK
DESK
DINNER

They all start with 'd'.

Now point to each picture and say its name after me:

DONKEY (pause), DRINK (pause), DESK (pause),
DINNER (pause).
They all start with 'd'.

Say:

Item 8

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

CHAIR
DUCK
DOLL
KITTEN

Now point to each picture and say its name after me:

CHAIR (pause), DUCK (pause), DOLL (pause),
KITTEN (pause).
Put a cross on the pictures that start with 'd'.
When all children have finished, say:

Preliminary exercise 9
NEST
Put your finger on each picture as I say its name:
NEEDLE
NUT
NOSE
They all start with 'n'.

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:
NEST...
NEEDLE...
NUT...
NOSE.

They all start with 'n'.

Item 9
SUN
NAIL
NET
PENCIL
They all start with 'n'.

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:
SUN...
NAIL...
NET...
PENCIL.

Now point to each picture and say its name after me:
SUN (pause), NAIL (pause), NET (pause),
PENCIL (pause).

Put a cross on the pictures that start with 'n'.

When all children have finished, say:

Preliminary exercise 10
GLASS
GOAT
GUITAR
GARDEN
They all start with 'g'.

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:
GLASS...
GOAT...
GUITAR...
GARDEN.

They all start with 'g'.

Item 10
GATE
KEY
ELEPHANT
GUN
They all start with 'g'.

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:
GATE...
KEY...
ELEPHANT...
GUN.

Now point to each picture and say its name after me:
GATE (pause), KEY (pause), ELEPHANT (pause),
GUN (pause).

Put a cross on the pictures that start with 'g'.

Remember, it's a secret.
Wait until all children have finished.
Say:

Turn the book over to page 5.
This is the page which has five stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.

Put your marker under the first row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:

SOAP
"SOAP ... STEPS ... SLIDE ... STAR."

They all start with 's'.

Now point to each picture and say its name after me:

SOAP (pause), STEPS (pause), SLIDE (pause),
STAR (pause).
They all start with 's'.

Now put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:

BOX
SPOON
"BOX ... SPOON ... SOCK ... COAT."

Now point to each picture and say its name after me:

BOX (pause), SPOON (pause), SOCK (pause),
COAT (pause).
Put a cross on the pictures that start with 's'.

When until all children have finished,
say:

Put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:

ROPE
RAKE
"ROPE ... RAKE ... RING ... RADIO."

They all start with 'r'.

(sound the letter, do not use the letter-name)

Now point to each picture and say its name after me:

ROPE (pause), RAKE (pause), RING (pause),
RADIO (pause).
They all start with 'r'.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981.
ItiM 12
HOOK
CAR
RABBIT
ROSE
I
Preliminary exercise 13
PANTS
PRAM
PAN
PIN
Item 12
Now put your marker under the next row of pictures, like this. (demonstrate)
HOOK
CAR
RABBIT
ROSE
Put your finger on each picture as I say its name:
HOOK...CAR...RABBIT...ROSE.
Now point to each picture and say its name after me:
HOOK (pause), CAR (pause), RABBIT (pause),
ROSE (pause).
Put a cross on the pictures that start with 'r'.
When all children have finished,
[p]
Preliminary
exercise 13
Put your marker under the next row of pictures, like this. (demonstrate)
PANTS
PRAM
PAN
PIN
Put your finger on each picture as I say its name:
PANTS...PRAM...PAN...PIN.
They all start with 'p'.
Now point to each picture and say its name after me:
PANTS (pause), PRAM (pause), PAN (pause),
PIN (pause).
They all start with 'p'.
Item 13
Now put your marker under the next row of pictures, like this. (demonstrate)
DRUM
POT
PIPE
CHAIR
Put your finger on each picture as I say its name:
DRUM...POT...PIPE...CHAIR.
Now point to each picture and say its name after me:
DRUM (pause), POT (pause), PIPE (pause),
CHAIR (pause).
Put a cross on the pictures that start with 'p'.
When all children have finished,
say:
Turn the page over to page 6.
Fold your book back.
This is the page which has six stars at the top.
Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Preliminary exercise 14

CROSS
CAT
COAT
KETTLE

Put your finger on each picture as I say its name:

CROSS ... CAT ... COAT ... KETTLE.

They all start with 'c'.

Now point to each picture and say its name after me:

CROSS (pause), CAT (pause), COAT (pause),
KETTLE (pause).

They all start with 'c'.

Item 14

COW
PENCIL
BUCKET
CART

Put your finger on each picture as I say its name:

COW ... PENCIL ... BUCKET ... CART.

Now point to each picture and say its name after me:

COW (pause), PENCIL (pause), BUCKET (pause),
CART (pause).

Put a cross on the pictures that start with 'c'.

Wait until all children have finished.

Preliminary exercise 15

HILLS
HOUSE
HEN
HAMMER

Put your finger on each picture as I say its name:

HILLS ... HOUSE ... HEN ... HAMMER.

They all start with 'h'.

Now point to each picture and say its name after me:

HILLS (pause), HOUSE (pause), HEN (pause),
HAMMER (pause).

They all start with 'h'.

Item 15

HAT
BED
HORSE
FEATHER

Put your finger on each picture as I say its name:

HAT ... BED ... HORSE ... FEATHER.

Now point to each picture and say its name after me:

HAT (pause), BED (pause), HORSE (pause),
FEATHER (pause).

Put a cross on the pictures that start with 'h'.
When all children have finished, say:
Put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
Vegetables ... Van ... Vest ... Visitor.
They all start with 'v'.
Now point to each picture and say its name after me:
Vegetables (pause), Van (pause), Vest (pause), Visitor (pause).
They all start with 'v'.
Now put your marker under the next row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
Crown ... Vacuum ... Bird ... Vase.
Now point to each picture and say its name after me:
Crown (pause), Vacuum (pause), Bird (pause), Vase (pause).
Put a cross on the pictures that start with 'v'.
When all children have finished, say:
Turn your book over to page 7.
This is the page which has seven stars at the top.
Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:
Put your marker under the first row of pictures, like this. (demonstrate)
Put your finger on each picture as I say its name:
Tin ... Tv ... Tennis ... Tiger.
They all start with 't'.
Now point to each picture and say its name after me:
Tin (pause), Tv (pause), Tennis (pause), Tiger (pause).
They all start with 't'.

ACER Early School Series. Recognition of Initial Consonant Sounds Test Directions for Administration. Copyright © ACER 1981
Item 17

Now put your marker under the next row of pictures, like this. (demonstrate)

Put your finger on each picture as I say its name:

**DOG**... **LOAF**... **TRUCK**... **TOOTHBRUSH**.

Now point to each picture and say its name after me:

**DOG** (pause), **LOAF** (pause), **TRUCK** (pause), **TOOTHBRUSH** (pause).

Put a cross on the pictures that start with 't'. Remember, it's a secret.

Wait until all children have finished.

Say:

**Now put down your crayons and close your books very quickly.**

Collect the test booklets as soon as the last item has been completed.

**Scoring** (Maximum score: 17)

If the test is administered in more than one session, leave the scoring until the whole test has been completed.

Score one point for each correct item. Possible points for each item are one (1) or zero (0).

Zero (0) is given for items in which either one or more wrong pictures, all pictures or no pictures have been crossed. For an item in which one correct and one incorrect picture are crossed, zero (0) points are given.

During scoring, record points in the box provided for each item in the right-hand margin. Record the total number of correct items where indicated on the front of the test booklet.

---

<table>
<thead>
<tr>
<th>Page</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Example</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>
Auditory Discrimination Test

Directions for Administration
Auditory Discrimination Test
Directions for Administration

Helga A.H. Rowe

Other items in the ACER Early School Series are:
Test Booklets
Score Keys
Early Identification and Intervention:
A Handbook for Teachers and School Counsellors

Copyright ©ACER 1981
Published by The Australian Council for Educational Research Limited, Radford House, Frederick Street, Hawthorn, Victoria, 3122.
EXPLANATORY NOTE

This test forms part of the ACER Early School Series, a set of diagnostic and screening instruments designed to provide teachers working with school beginners with a variety of approaches for the early identification of certain abilities and developmental needs of individual children.

All tests in the series are specifically designed to be used by teachers. This decision is based on the premise that the classroom teacher plays the key role in the early identification of children who may be at risk for learning handicaps or other school problems. It is hoped that these tests may complement other ways of assessment and the teacher's judgment, in identifying children who may require more individual attention to develop their learning potential.

The planning of classroom activity needs to take into consideration not only the individual differences between children in one beginners' grade, but also the considerable differences which are frequently found within one and the same child with respect to the development of language and other skills, personal experience, knowledge, and other resources.

The ACER Early School Series is aimed at identifying levels the child has reached by providing a standardized assessment through tasks which are based on regular classroom activity and which, through minimal interruption to the teaching program, may aid in adjusting instruction to the individual child's abilities and needs.

This booklet contains general and detailed directions for the administration of the Auditory Discrimination Test.

Other tests contained in the series are:

- Figure Formation Test
- Number Test
- Recognition of Initial-Consonant Sounds Test
- Word Knowledge Test
- Receptive Language Skills
  - Comprehension Test
  - Negation Test
  - Prepositions Test
  - Pronouns Test
  - Verb Tense Test

Early Identification and Intervention: A Handbook for Teachers and School Counsellors is common to all 10 tests, and contains information concerning the rationale of the tests, their use and interpretation, and details pertaining to their development.

*ACER Early School Series, Auditory Discrimination Test Directions for Administration. Copyright © ACER 1981.*
GENERAL DIRECTIONS FOR
ADMINISTRATION

Advance Planning
To ensure that the results of testing are as valid and reliable as possible, it is suggested that teachers familiarize themselves with the following instructions during the planning stage some time before the day of testing:

1. Familiarize yourself with *Early Identification and Intervention: A Handbook for Teachers and School Counsellors*, with the test booklets, and with the Detailed Directions for Administration contained in this booklet.
2. Ensure that sufficient test booklets are available, including a few extra copies for emergencies.
3. Prior to the day of testing, enter the child's name and other details required on the front of each test booklet.
4. Have available a sufficient supply of crayons or pencils and markers for each child.
5. Have available a box of tissues.
6. Children should be tested in small groups of no more than eight children. If at all possible, no more than five children should be tested at once.
7. Attempt to obtain the help of another teacher as assistant for the days of testing.
8. Arrange place for testing.
9. Ensure that you are not disturbed while you are administering a test.

On the Day of Testing
1. Have this booklet open at the pages giving the detailed directions.
2. Provide yourself with an unused test booklet for the test you are about to administer. This is essential as you are required to refer to the test booklet to demonstrate during the reading of the detailed instructions.
   A useful technique for the administration of the tests may be to hold the test booklet open at the appropriate page so that it faces the children. Place your copy of the Directions in such a way that you can read the directions and at the same time point to the appropriate section on the test booklet, so that the children can see exactly what you are referring to.
3. Have available paper and pencil for yourself in case it is necessary for you to make notes.
4. See that the tables or desks used during the testing are cleared.
5. Arrange seating in such a way as to avoid copying.
6 Arrange for all children to be tested to visit the toilet just before the start of the test.
7 Lay out a test booklet, crayon and marker for each child.
8 Check the names on each test booklet to ensure that each child has the correct booklet.

During Testing
1 Make every effort to maintain a manner and classroom atmosphere which are as natural as possible.
2 Minimize distraction during the test.
3 Read the directions for each item exactly as printed.
4 During the administration of practice and example items (marked by letters in the directions) spend as much time as necessary to make sure that each child understands what is required. It is important that each child approaches the test items with confidence.
5 In the case of test items (numbered items) do not give any help beyond reading and demonstrating the instructions as printed.

Scoring — General Rules
1 Do not try to score the test without using the Score Key. The Key is provided to make the scoring objective, quick, and accurate.
2 Example items and preliminary exercises are not to be scored.
3 Record the number of points scored in the margin next to each item. After the test has been marked, place the total score after ‘Number of correct items’ on the front of each test booklet.
4 All scoring should be re-checked to ensure maximum accuracy.
5 Follow the specific directions given for scoring this test.
6 Much time is saved and the likelihood of errors is reduced if a single test is scored for the entire class at one time.
DETAILED DIRECTIONS FOR ADMINISTRATION

Materials required: Each child should have a test booklet, a cardboard marker (size 12.5 cm x 20 cm, 5" x 8") and a crayon or pencil.

To administer the test follow the directions exactly as given here. The test may be given in one session or broken up and given in two or three sessions. If the test is to be administered in sections over a number of days, administer the three example items (A, B, C) at the beginning of each testing session and then proceed to the item following the one which was administered last.

Do not elaborate on the pictures of the test items. Make sure that there is no interfering noise. Speak clearly and distinctly. Do not repeat any stimulus word unless you have made an error.

Place a test booklet in front of each child so that the front page is face up. Make sure that each child has the necessary test materials. Wait until all children are ready to begin.

Say:
Today we are going to work with pictures.
In each row there are two pictures to choose from.
You must listen very carefully as I say the name of the picture I want you to put a cross on.

Example A
MOUSE
Put your marker under the first row of pictures like this. (demonstrate)

Here we have a picture of a HOUSE and a picture of a MOUSE.
Find the MOUSE.
With your crayon put a cross on the picture of the MOUSE.
(demonstrate)
Make sure that each child has put a cross on the correct picture. Assist children who have not put a cross on the picture of the MOUSE.

**Example B**

**COOK**

Say:

Put your marker under the next row of pictures.
(demonstrate)

Here is a picture of a BOOK and a picture of a COOK. Find the COOK. Put a cross on the picture of the COOK.

Assist children who have not put a cross on the correct picture.

**Example C**

**HAT**

Say:

Put your marker under the next row of pictures.
Here we have a picture of a HAT and a picture of a MAT. Put a cross on the HAT.

Assist children who have not put a cross on the correct picture. Repeat all example items until you believe each child clearly understands the instructions.

Say:

Turn the page over to page 2. Fold your book back. This is the page which has two stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:
Put your marker under the top row.
Notice the triangle (Δ) at the beginning and at the end of the row. (pause) KEY, BEE.
Find the BEE. Put a cross on the BEE.

Wait until all children have finished.

Say:
Put your marker under the next row.
Notice the star (*) at the beginning and at the end of the row. (pause) SOAP, ROPE.
Find the SOAP. Put a cross on the SOAP.

Check that markers are in the correct position.

When all children have finished, say:
Put your marker under the next row.
Notice the square (□) at the beginning and at the end of the row. (pause) BED, BELL.
Find BED. Put a cross on BED.

When all children have finished, say:
Put your marker under the next row.
Notice that there is a dot (●) at the beginning and at the end of the row.
Place your marker in such a way that you can see two dots and two pictures in that row. (pause) CAT, BAT.
Put a cross on the BAT.
When all children have finished, say:

**Item 5**

+ SAUCE +

Notice the cross (+) at the beginning and at the end of the row. (pause) HORSE, SAUCE.

Put a cross on SAUCE.

When all children have finished, say:

**Item 6**

*PUZZLE*

Notice the star (*) at the beginning and at the end of the row. (pause) PUDDLE, PUZZLE.

Put a cross on PUZZLE.

When all children have finished, say:

Turn the book over to page 3.

This is the page which has three stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.

**Item 7**

○ TREE ○

Notice the circle (○) at the beginning and at the end of the row. (pause) TRAY, TREE.

Find TREE. Put a cross on TREE.

Wait until all children have finished.

Say:

**Item 8**

△ SHIP△

Notice the triangle (△) at the beginning and at the end of the row. (pause) SHIP, SHOP.

Find SHIP. Put a cross on SHIP.
When all children have finished,
say:

**Item 9**

Put your marker under the next row.
*FEATHER* Notice the dot (•) at the beginning and at the end of the row. (pause) FEATHER, FATHER.
Put a cross on FEATHER.

When all children have finished,
say:

**Item 10**

□SEE SAW□ Notice the square (□) at the beginning and at the end of the row. (pause) SEE SAW, SEA SHORE.
Put a cross on SEE SAW.

When all children have finished,
say:

**Item 11**

•FISH• Notice the star (*) at the beginning and at the end of the row. (pause) FISH, DISH.
Put a cross on the FISH.

When all children have finished,
say:

**Item 12**

△ BELL △ Notice the triangle (△) at the beginning and at the end of the row. (pause) BELL, SHELL.
Put a cross on the BELL.

When all children have finished,
say:

Turn the page over to page 4. Fold your book back.
This is the page which has four stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.
Say:

**Item 13**

Put your marker under the top row.

○ CIRCLE ○

Notice the circle (○) at the beginning and at the end of the row. (pause) CLOCK, BLOCK.

Find the CLOCK. Put a cross on the CLOCK.

When all children have finished,
say:

**Item 14**

Put your marker under the next row.

□ MONEY □

Notice the square (□) at the beginning and at the end of the row. (pause) BUNNY, MONEY.

Put a cross on MONEY.

When all children have finished,
say:

**Item 15**

Put your marker under the next row.

△ ROCK, △

Notice the triangle (△) at the beginning and at the end of the row. (pause) ROCK, SOCK.

Put a cross on the ROCK.

When all children have finished,
say:

**Item 16**

Put your marker under the next row.

• PLANE •

Notice the dot (•) at the beginning and at the end of the row. (pause) PLATE, PLANE.

Put a cross on the PLANE.

When all children have finished,
say:

**Item 17**

Put your marker under the next row.

+ GRASS +

Notice the cross (+) at the beginning and at the end of the row. (pause) GLASS, GRASS.

Put a cross on GRASS.
When all children have finished, say:
Item 18
Put your marker under the next row.
Notice the star (*) at the beginning and at the end of the row. (pause) KITTEN, COTTON.
Put a cross on COTTON.

When all children have finished, say:
Turn your book over to page 5. This page has five stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.

Say:
Item 19
Put your marker under the top row.
Notice the triangle (△) at the beginning and at the end of the row. (pause) TOOL, TOOTH.
Find the TOOTH. Put a cross on the TOOTH.

When all children have finished, say:
Item 20
Put your marker under the next row.
Notice the cross (+) at the beginning and at the end of the row. (pause) PRA/PLAY.
Put a cross on PLAY.

When all children have finished, say:
Item 21
Put your marker under the next row.
Notice the triangle (△) at the beginning and at the end of the row. (pause) PIG, PIN.
Put a cross on PIG.

ACER Early School Series. Auditory Discrimination Test Directions for Administration. Copyright © ACER 1981
When all children have finished, say:

Item 22

Put your marker under the next row.
Notice the circle (o) at the beginning and at the end of the row. (pause) SLEEP, SWEEP.
Find SWEEP. Put a cross on SWEEP.

When all children have finished, say:

Item 23

Put your marker under the next row.
Notice the cross (+) at the beginning and at the end of the row. (pause) SHEEP, SHIP.
Put a cross on SHEEP.

When all children have finished, say:

Item 24

Put your marker under the next row.
Notice the square (□) at the beginning and at the end of the row. (pause) TOE, TWO.
Put a cross on TOE.

When all children have finished, say:

Item 25

Put your marker under the top row.
Notice the triangle (△) at the beginning and at the end of the row. (pause) PEN, HEN.
Find the PEN. Put a cross on the PEN.

ACER Early School Series. Auditory Discrimination Test Directions for Administration. Copyright © ACER 1981
When all children have finished,
say:

**Item 26**  *WASH*

Put your marker under the next row.
Notice the star (*) at the beginning and at the end of the row. (pause) *WASH, WASH.*
Put a cross on WASH.

When all children have finished,
say:

**Item 27**  + ROOF +

Put your marker under the next row.
Notice the cross (+) at the beginning and at the end of the row. (pause) *ROOM, ROOF.*
Put a cross on ROOF.

When all children have finished,
say:

**Item 28**  ØPETØ

Put your marker under the next row.
Notice the circle (Ø) at the beginning and at the end of the row. (pause) *PET, PEN.*
Find the PET. Put a cross on PET.

When all children have finished,
say:

**Item 29**  □FRUIT□

Put your marker under the next row.
Notice the square (□) at the beginning and at the end of the row. (pause) *FRUIT, FLUTE.*
Put a cross on FRUIT.

When all children have finished,
say:

**Item 30**  △STOOL△

Put your marker under the next row.
Notice the triangle (△) at the beginning and at the end of the row. (pause) *SCHOOL, STOOL.*
Find STOOL. Put a cross on STOOL.

When all children have finished,
say:

Turn your book over to page 7. This page has seven stars at the top.
Point to the stars at the top of the page. Wait until all children have turned to the correct page.

ACER Early School Series. Auditory Discrimination Test. Directions for Administration. Copyright © ACER 1981
Say:  
**Item 31**  
* + PIN +  
* Notice the cross (+) at the beginning and at the end of the row. (pause) PIN, PEN.  
Find PIN. Put a cross on PIN. 

When all children have finished, say:  
**Item 32**  
* PATH*  
* Notice the square (□) at the beginning and at the end of the row. (pause) PARK, PATH.  
Put a cross on the PATH.  

Wait until all children have finished.  
Say:  
**Item 33**  
* TOWN*  
* Notice the dot (●) at the beginning and at the end of the row. (pause) TOWN, TOWEL.  
Put a cross on TOWN.  

When all children have finished, say:  
**Item 34**  
* + PEARL +  
* Notice the cross (+) at the beginning and at the end of the row. (pause) GIRL, PEARL.  
Put a cross on PEARL.  

Wait until all children have finished.  
Say:  
**Item 35**  
* PATH*  
* Notice the square (□) at the beginning and at the end of the row. (pause) MOUTH, MOUSE.  
Put a cross on MOUTH.  

When all children have finished, say:  
**Item 36**  
* △ PEEL △  
* Notice the triangle (△) at the beginning and at the end of the row. (pause) POOL, PEEL.  
Put a cross on PEEL.  

ACER Early School Series. Auditory Discrimination Test. Directions for Administration. Copyright © ACER 1981
When all children have finished, say:

Turn the page over to page 8. Fold your book back. This page has eight stars at the top.

Point to the stars at the top of the page. Wait until all children have turned to the correct page.

Item 37. Put your marker under the first row.

KNOT. Notice the dot (●) at the beginning and at the end of the row. (pause) KNOT, NUT.

Find KNOT. Put a cross on KNOT.

When all children have finished, say:

Item 38. Put your marker under the next row.

BOARD. Notice the cross (+) at the beginning and at the end of the row. (pause) BOARD, BALL.

Put a cross on the BOARD.

When all children have finished, say:

Item 39. Put your marker under the next row.

CAP. Notice the circle (○) at the beginning and at the end of the row. (pause) CAP, CAT.

Put a cross on the CAP.

When all children have finished, say:

Item 40. Put your marker under the next row.

FOUNTAIN. Notice the star (*) at the beginning and at the end of the row. (pause) FOUNTAIN, MOUNTAIN.

Put a cross on the FOUNTAIN.

When all children have finished, say:

Now put down your crayons and close your books very quickly.

Collect the booklets as soon as the last item has been completed.

**Scoring (Maximum score: 40)**

If the test is administered in more than one session, leave the scoring until the whole test is completed.

Do not score example items.

Score one point for each correct item crossed, i.e. possible points for each item are zero (0), or one (1). Zero (0) is given for items in which either the wrong pictures, both pictures, or neither of the pictures has been crossed.

During scoring, record points in the right-hand margin next to each item. Record the total number of correct items where indicated on the front of the test booklet.

<table>
<thead>
<tr>
<th>Page</th>
<th>Item</th>
<th>Page</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ex. A house</td>
<td>mouse</td>
<td>5</td>
</tr>
<tr>
<td>Ex. C hat</td>
<td>hat</td>
<td>mat</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>key</td>
<td>bee</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>soap</td>
<td>rope</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>cat</td>
<td>bell</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>horse</td>
<td>sauce</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>puddle</td>
<td>puzzle</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>tray</td>
<td>tree</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>ship</td>
<td>shop</td>
<td>29</td>
</tr>
<tr>
<td>9</td>
<td>feather</td>
<td>father</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>see-saw</td>
<td>sea shore</td>
<td>31</td>
</tr>
<tr>
<td>11</td>
<td>fish</td>
<td>dish</td>
<td>32</td>
</tr>
<tr>
<td>12</td>
<td>bell</td>
<td>shell</td>
<td>33</td>
</tr>
<tr>
<td>13</td>
<td>clock</td>
<td>block</td>
<td>34</td>
</tr>
<tr>
<td>14</td>
<td>bunny</td>
<td>money</td>
<td>35</td>
</tr>
<tr>
<td>15</td>
<td>rock</td>
<td>sock</td>
<td>36</td>
</tr>
<tr>
<td>16</td>
<td>plate</td>
<td>plane</td>
<td>37</td>
</tr>
<tr>
<td>17</td>
<td>glass</td>
<td>grass</td>
<td>38</td>
</tr>
<tr>
<td>18</td>
<td>kitten</td>
<td>cotton</td>
<td>39</td>
</tr>
<tr>
<td>19</td>
<td>tool</td>
<td>tool</td>
<td>40</td>
</tr>
</tbody>
</table>

ACER
Early
School
Series
Pronouns Test

Name: 
Date of Birth: 
Home Language: 
Comments: 

Sex: 
Today's Date: 
Class: 
Number of Correct Items: 

A
ACER Early School Series

Figure Formation Test

Name: ____________________________ Sex: ________ School: ________________________________

Date of Birth: ____________________________ Today's Date: __________________________

Home Language: ____________________________ Number of Correct Items: __________

Comments: ____________________________
ACER Early School Series

Number Test

Name: 
Date of Birth: 
Home Language: 
Comments: 

Sex: 
School: 
Class: 

Today's Date: 
Number of Correct Items: 

299
ACER
Early
School
Series
Recognition of Initial
Consonant Sounds Test

Name:
Date of Birth:
Sex:
Today's Date:
School:
Class:
Home Language:
Number of Correct Items:
Comments:

[Grid of nine images representing sounds: P, Ex, bone, tent, bag, Australia, telephone, bird, dog, bicycle]
ACER Early School Series

Auditory Discrimination Test

Name: [Blank]
Date of Birth: [Blank]
Sex: [Blank]
Today's Date: [Blank]
School: [Blank]
Class: [Blank]
Home Language: [Blank]
Number of Correct Items: [Blank]
Comments: [Blank]

ExA

ExB

ExC

ExD

Today's Date: [Blank]
Sex: [Blank]
School: [Blank]
Class: [Blank]
Home Language: [Blank]
Number of Correct Items: [Blank]
Comments: [Blank]

ExA

ExB

ExC

ExD
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td><img src="image" alt="Clock" /></td>
<td><img src="image" alt="Speakers" /></td>
</tr>
<tr>
<td>14</td>
<td><img src="image" alt="Rabbit" /></td>
<td><img src="image" alt="Money" /></td>
</tr>
<tr>
<td>15</td>
<td><img src="image" alt="Rock" /></td>
<td><img src="image" alt="Chair" /></td>
</tr>
<tr>
<td>16</td>
<td><img src="image" alt="Plate" /></td>
<td><img src="image" alt="Airplane" /></td>
</tr>
<tr>
<td>17</td>
<td><img src="image" alt="Glass" /></td>
<td><img src="image" alt="Grass" /></td>
</tr>
<tr>
<td>18</td>
<td><img src="image" alt="Cat" /></td>
<td><img src="image" alt="Thread" /></td>
</tr>
</tbody>
</table>
Word Knowledge Test

Score Key
ACER
Early
School
Series

Comprehension Test

Score Key
ACER
Early School Series

Negation Test

Score Key
Prepositions Test

Score Key
ACER
Early School Series

Figure Formation Test

Score Key
ACER
Early
School
Series

Number Test

Score Key
1

3

4

372

373
Recognition of initial Consonant Sounds Test

Score Key