

<<语言测试/牛津语言学入门丛书>>

图书基本信息

书名：<<语言测试/牛津语言学入门丛书>>

13位ISBN编号：9787810806589

10位ISBN编号：7810806580

出版时间：2003-2

出版时间：上海外语教育出版社

作者：麦克纳马拉，，，澳大利亚 编

页数：140

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

内容概要

《牛津语言学入门丛书：语言测试》是一部有关语言测试基本理论和实践的导论性著作，内容涵盖了测试设计、测试评分程序、测试的效度和信度以及语言测试的社会因素等方面。

书籍目录

Preface Author's preface SECTION I Survey 1 Testing, testing... What is a language test? Understanding language testing Types of test Test purpose The criterion The test-criterion relationship Conclusion 2 Communication and the design of language tests Discrete point tests Integrative and pragmatic tests Communicative language tests Models of communicative ability Conclusion 3 The testing cycle Understanding the constraints Test content Test method Authenticity of response Fixed and constructed response formats Test specifications Test trials Conclusion 4 The rating process Establishing a rating procedure The problem with raters Establishing a framework for making judgements Rating scales Holistic and analytic ratings Rater training Conclusion 5 Validity : testing the test Threats to test validity Test content Test method and test construct The impact of tests Conclusion 6 Measurement Introduction Measurement Quality control for raters Investigating the properties of individual test items Norm-referenced and criterion-referenced measurement New approaches to measurement Conclusion 7 The social character of language tests Introduction The institutional character of assessment Assessment and social policy Assessment and educational policy The social responsibility of the language tester Ethical language testing Accountability Washback Test impact Codes of professional ethics for language testers Critical language testing Conclusion 8 New directions-and dilemmas?

章节摘录

An important aspect of a scale is the way in which performance at the top end of the scale is defined. There is frequently an unacknowledged problem here. Rating scales often make reference to what are assumed to be the typical performances of native speakers or expert-users of the language at the top end of the scale. That is, it is assumed that the performance of native speakers will be fundamentally unlike the performances of non-native speakers, who will tend gradually to approximate native speaker performance as their own proficiency increases. However, claims about the uniformly superior performance of these idealized native speakers have rarely been supported empirically. In fact, the studies that have been carried out typically show the performance of native speakers as highly variable, related to educational level, and covering a range of positions on the scale. In spite of this, the idealized view of native speaker performance still hovers inappropriately at the top of many rating scales. The number of levels on a rating scale is also an important matter to consider, although the questions raised here are more a matter of practical utility than of theoretical validity. There is no point in proliferating descriptions outside the range of ability of interest. Having too few distinctions within the range of such ability is also frustrating, and the revision of rating scales often involves the creation of more distinctions. The failure of rating scales to make distinctions sufficiently fine to capture progress being made by students is a frequent problem. It arises because the purposes of users of a single assessment instrument may be at odds. Teachers have continuous exposure to their students' achievements in the normal course of learning. In the process, they receive ongoing informal confirmation of learner progress which may not be adequately reflected in a category difference as described by a scale. Imagine handing parents who are seeking evidence of their child's growth a measuring stick with marks on it only a foot (30 centimetres) apart, the measure not allowing any other distinction to be made. The parents can observe the growth of the child: they have independent evidence in the comments of relatives, or the fact that the child has grown out of a set of clothes. Yet in terms of the measuring stick no growth can be recorded because the child has not passed the magic cut-point into the next adjacent category of measurement. Teachers restricted to reporting achievement only in terms of broad rating scale categories are in a similar position. Most rating scales used in public educational settings are imposed by government authorities for purposes of administrative efficiency and financial accountability, for which fine-grained distinctions are unnecessary. The scales are used to report the achievements of the educational system in terms of changes in the proficiency of large numbers of learners over relatively extended periods of time. The government needs the 'big picture' of learner (and teacher) achievement in order to satisfy itself that its educational budget is yielding results. Teachers working with these government imposed, scale-based reporting mechanisms experience frustrations with the lack of fine distinctions on the scale. The coarse-grained character of the categories may hardly do justice to the teachers' sense of the growth and learning that has been achieved in a course. The purposes of the two groups—administrators, who are interested in financial accountability, and teachers, who are interested in the learning process may be at odds in such a case.

版权说明

本站所提供下载的PDF图书仅提供预览和简介, 请支持正版图书。

更多资源请访问:<http://www.tushu007.com>