# 中外语言测试基本理论与实践

修旭东　著

青岛海洋大学出版社
·青　岛·

**图书在版编目(CIP)数据**

**中外语言测试基本理论与实践/修旭东著. —青岛:青岛海洋大学出版社,**2002.9

ISBN 7-81067-388-2

Ⅰ.**中**... Ⅱ.**修**... Ⅲ.**语言能力—测试—研究** Ⅳ.H09

**中国版本图书馆** CIP **数据核字**(2002)**第** 066206 **号**

**青岛海洋大学出版社出版发行**

(**青岛市鱼山路** 5 **号** **邮政编码**:266003)

**出版人**:**李学伦**

**日照日报社印刷厂印刷**

**新华书店经销**

*

**开本**:850mm×1 168mm 1/32 **印张**:9.0 **字数**:220 **千字**

2002 **年** 10 **月第** 1 **版** 2002 **年** 10 **月第** 1 **次印刷**

**印数**:1～3 000 **册** **定价**:17.60 **元**

# FOREWORD

This book on testing is valuable because there is a pressing need in all sections of the education community for a greater awareness of the test development process. This book aims to raise the level of professional competence in developing reliable and valid tests. We need to develop tests that actually test what we want to test and which provide us with results that we can depend on, i. e., our tests must be both valid and reliable. The volume covers an impressive range of the areas we need to understand in language testing and will help readers develop their knowledge and skills in the area.

It has been said the language testing is the cutting edge of Applied Linguistics. It is through language testing that we will gain insight into the nature of language proficiency and also establish more clearly the development of language acquisition. Testing is not just an academic pursuit, however as its importance in the learning cycle shows.

Teachers need information on student progress during the life of a course as well as summative assessments

of ability at the end. Well constructed tests can provide this information. Teachers need to evaluate formatively to make appropriate decisions concerning modifications to teaching procedures and learning activities; to steer their way through the syllabus in action, i.e., to shape and influence the process. Is new material being introduced too quickly or too slowly? If it is, the effectiveness of learning may well be diminished. Teachers need to decide when to move on in a unit or to the next unit. If the next units are dependant on what has gone before then the teacher needs to be sure the students have mastered the objectives of a particular unit before proceeding. This can lead to necessary modifications in a programme.

There is a need to monitor the performance ability of students; a need to establish as clearly as possible what it is they can do. This can lead to a descriptive profile of a learner's communicative performance or a record of achievement. Formal tests devised for monitoring achievement can be extremely helpful to individual students, can help identify areas of strength or weakness and provide focus for future learning objectives. They can also be motivational by providing an idea of progress.

Tests have a valuable formative and summative role in education. It is the aim of this book to help practitioners in China to play their part.

**C. J. Weir**
**Aug. 2002**

# 前　言

学好、用好语言是人类文明与进步的必要条件之一。从文化人类学的角度来看，由于种族和民族的不同，语言可以有母语、第二语言和外语之分。但是，从语言学习与应用的角度来看，无论是母语还是母语之外的其他语言，都存在着一个怎样才算是学好和用好语言的问题。那么，怎样才算学好、用好了语言呢？从事物的相对性和绝对性来看，“学好、用好”只能是相对的，学无止境才是绝对的。尽管如此，要看是否真正地“学好、用好”，总需要一定的标准来衡量。需要什么样的标准、怎样来测量，这些都需要相关的理论和依据作指导。改革开放以后，以英语语言测试研究为代表的理论与实践在我国产生了广泛而深远的影响，这对我国的英语教学与评估是一件非常有意义的事情，而在语言共性的基础上，它对汉语和其他语种的教学与评估，无疑也会产生有益的影响。事实上，在我国，不仅仅语言测试，甚至其他科目的测试乃至对教育质量的评估指标，均需进行深入的研究，否则，素质教育将难以全面而科学地进行。实践中，尽管强调不要把分数当成“指挥棒”，但很少有人能够真正做到。在这种情况下，测试的质量就显得尤为重要。多数人会认为，得 90 分的考生比得 85 分的考生优秀，但大多数人所采用的这种判断

“标准”在多大程度上具有科学性呢？这就很难回答了。我们或许可以回答，到目前为止某某考试是比较科学或合乎标准的，但永远不可能回答某某考试已经是最科学、最合乎标准的。测试的质量如同商品的质量一样，“没有最好，只有更好”。如果“最好”了，为什么高考后会出现那么多“高分低能”的人呢？Hughes(1989)认为：事实上，没有最好的测试或测试技术。既然“没有最好、只有更好”，对语言测试同样需要永不间断地进行探索与研究。

本书所论述的“中外语言测试”是指国内外对以英语为代表的语言所进行的测试，作者在该书中主要对国内外以英语为实例的语言测试理论和实践进行阐述和研究。

从语言学角度看，语言测试属于应用语言学研究的范畴。同时，从测试学角度看，它要符合测试的基本要求。“测试”的定义有许多种，目前公认较为权威的是 Anastasi(1982)的定义(刘润清、韩宝成，2000)：测试是对行为样本所做的客观的标准化的测量。该定义包含三个基本因素：(1) 行为样本：在语言测试中，它是指对语言能力表现行为的有效的抽样。(2) 客观的测量：在语言测试中，它主要指测试项目的难易度和区分度，测试结果的可靠性和有效性。(3) 标准化的测量：在语言测试中，它主要指测试题目的编制、测试的实施、分数的解释等一整套系统程序所具备的严密性，从而使测试结果具有可比性。

Spolsky(1979)把语言测试的发展分为三个阶段：前

科学阶段(Pre-scientific trend)、心理测量与结构主义阶段(Psychometric-structuralist trend)和心理学与社会语言学阶段。

根据测试的定义和语言测试发展阶段的划分,本书主要介绍和研究了语言测试第三个阶段的理论与实践。在这一阶段中,我们可以看到:Bachman(1990)关于语言测试理论(信度、效度、项目分析等)的阐述成为语言测试界的里程碑;Weir(1993)关于四项基本技能的测试框架为语言测试的设计提供了蓝图;桂诗春教授对分数统计与解析的阐释开了我国英语标准化测试的先河。本书的特点是把国内外语言测试的主要理论与实践进行了比较全面、系统和由浅入深的介绍和探讨。

本书的主要内容包括:语言测试的主要类型、原则(信度、效度、项目分析等)、试卷的总体设计、四项基本技能的测试框架、分数的分析与解释等。对语言测试的前沿进行探讨的章节主要有:信度的综合归纳理论、项目反映理论、四项基本技能测试的真实性和交互性原则。

本书的适用对象包括英语本科高年级学生、语言学研究生、外语教学工作者等,并对从事汉语及其他语种教学的工作者也有一定的参考价值。

本书参考了大量国内外原著,在成书过程中,英国的Don Porter 教授、Cyril J. Weir 教授给予了热情的指导与帮助,Weir 教授在百忙中为本书作序,曾与本人同在英国里丁大学语言科学系作过访问学者的朋友以及在国内

工作的亲友们对本书也提出过许多宝贵意见，烟台师范学院外语系的有关同事也给予了许多帮助，在此一并致谢！

由于笔者水平所限，书中错误、缺点难免，敬请指正。

修旭东

2002 年 6 月

# 目　录

# Contents