

..... 剑·桥·英·语·教·师·丛·书

Assessing Reading.....

# 阅读评价

J. Charles Alderson (英) 著



CAMBRIDGE

E n g l i s h T e a c h e r

..... 剑 · 桥 · 英 · 语 · 教 · 师 · 丛 · 书

Assessing Reading.....

# 阅读评价

J. Charles Alderson (英) 著  
王笃勤 导读

外语教学与研究出版社 北京

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS BEIJING

剑桥大学出版社

CAMBRIDGE UNIVERSITY PRESS

京权图字：01-2010-6795

This is a reprint edition of the following title published by Cambridge University Press:

Assessing Reading (ISBN: 9780521599993)

© Cambridge University Press 2000

This reprint edition for the People's Republic of China (excluding Hong Kong SAR, Macao SAR and Taiwan Province) is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Foreign Language Teaching and Research Press and Cambridge University Press 2011

This reprint edition is authorized for sale in the People's Republic of China (excluding Hong Kong SAR, Macao SAR and Taiwan Province) only. Unauthorized export of this reprint edition is a violation of the Copyright Act. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of both Cambridge University Press and Foreign Language Teaching and Research Press.

本书版权由剑桥大学出版社和外语教学与研究出版社共同所有。本书任何部分之文字及图片，如未获得两社书面同意，不得用任何方式抄袭、节录或翻印。

只限中华人民共和国境内销售，不包括香港特别行政区、澳门特别行政区及台湾省。不得出口。

### 图书在版编目(CIP)数据

阅读评价 = Assessing Reading: 英文 / (英) 奥尔德森 (Alderson, J. C.) 著; 王笃勤导读. — 北京: 外语教学与研究出版社, 2011.8  
(剑桥英语教师丛书)  
ISBN 978-7-5135-1127-8

I. ①阅… II. ①奥… ②王… III. ①英语—阅读教学—教育评估—英文 IV. ①H319.4

中国版本图书馆 CIP 数据核字 (2011) 第 163457 号



悠游网—外语学习 一网打尽

www.2u4u.com.cn

阅读、视听、测试、交流、共享

提供海量电子文档、视频、MP3、手机应用下载!

出版人：蔡剑峰

项目负责：范海祥

责任编辑：范海祥

执行编辑：郑加丽

封面设计：张 峰

出版发行：外语教学与研究出版社

社 址：北京市西三环北路 19 号 (100089)

网 址：<http://www.fltrp.com>

印 刷：三河市北燕印装有限公司

开 本：650×980 1/16

印 张：29

版 次：2011 年 8 月第 1 版 2011 年 8 月第 1 次印刷

书 号：ISBN 978-7-5135-1127-8

定 价：79.00 元

\*

\*

\*

购书咨询：(010)88819929 电子邮箱：[club@fltrp.com](mailto:club@fltrp.com)

如有印刷、装订质量问题，请与出版社联系

联系电话：(010)61207896 电子邮箱：[zhijian@fltrp.com](mailto:zhijian@fltrp.com)

制售盗版必究 举报查实奖励

版权保护办公室举报电话：(010)88817519

物料号：211270001

# 出版说明

随着国家《英语课程标准》的实施以及《国家中长期教育改革和发  
展规划纲要》(2010—2020年)的颁布,英语教育领域正在进行着全面  
而深入的改革。新课程改革要求教师提高素质、更新观念、转变角色,  
同时要求教师的教学行为发生相应的变化。研究学生如何学、教师如何  
教、考试如何考和教师如何在教学实践中提高自我素质已经成为目前英  
语语言教学研究的焦点。因此,教师培训工作的重要性和关键作用变得  
越来越突出。为了更好地配合新课程改革,提高教师的自我素质,外语  
教学与研究出版社精选并引进了剑桥大学出版社设计的这套“剑桥英语  
教师丛书”。

这套丛书由当今国际语言教育界的权威人士 Gary Buck, John Read,  
J. Charles Alderson 等编著,由国内英语教育界专家、学者撰写导读。这  
些专家、学者在认真研读原著的基础上,结合中国外语教师的需求和现  
状,有针对性地用深入浅出的语言对原书作了概括、总结与评价,有助  
于英语教师及英语专业的研究生对原著的学习与理解。

本丛书的选题都是目前英语教师关注的教学实践中的热点专题。有  
几本是关于英语教学中的测试方面的,如 *Assessing Vocabulary, Assessing  
Listening, Assessing Speaking, Assessing Reading, Assessing Writing*; 有  
关于语言教师开展自我研究的,如 *Psychology for Language Teachers*; 有  
关于将信息技术用于语言教学之中的,如 *Language Learning in Distance  
Education, Assessing Language through Computer Technology*。

本丛书从英语语言教学的理论和实践两个层面为英语教师的教学实践提供理论指导，并为英语教师的在职教育和终生职业发展提供丰富的资源，可以作为英语教师专业化教育与发展和英语专业研究生教育的教材，供在职英语教师和英语语言教学研究者，尤其是英语专业的研究生学习使用。

外语教学与研究出版社

# 阅

## 读评价

### 导 读

尽管学术界存在各种不同的评价理论，但是，评价的价值判断本质已经为专家和学者们所公认，他们普遍认为评价是一种价值判断的过程（Stufflebeam, 1974；冯平，1995），是判断价值客体属性满足价值主体需求程度的一个过程。评价研究者提出了各种评价模式，如：形成性与终结性评价、直接与间接评价、分离式与综合性评价等，表现性评价、真实性评价、动态评价、发展性评价也受到越来越多的关注。其中，目前被广泛采纳的依旧是形成性评价与终结性评价，各级课程、大纲中对评价的描述也多采用形成性与终结性的表述方式。

阅读是日常生活和工作中最常用的技能，是人们获取信息和学习知识最主要的手段，阅读评价更是多种类型考试的重要组成部分。但是，阅读评价自从有评价之日起就没有发生过太多的变化，评价的内容和方式趋同，评价结果报告的形式不变，不能为测试使用者提供应有的阅读能力解释。阅读评价关注更多的是读者的阅读结果，或者评价者预期的阅读结果，很少关注阅读过程，因此很难为阅读教学提供有益

的参考。本书回顾了大量阅读测试与评价方面的相关研究，从阅读过程和阅读结果两个方面分析了目前阅读评价和测试中存在的问题，并结合目标语使用语境下的真实阅读探讨了阅读过程和阅读结果的评价方式，是语言测试从业者、研究者，甚至一线教师不可或缺的学术专著。

# 1

## 全书结构和总体评价

---



### 全书结构

本书共由九章组成。

第一章为“阅读的本质”，探索了阅读的过程与结果、理解的层次、阅读能力的内涵、阅读的模式、阅读与认知的关系、母语阅读与二语阅读等。

第二章为“影响阅读本质的因素”，主要从读者因素和文本因素两方面介绍了影响阅读的因素。

第三章为“阅读评价研究”，探讨了影响阅读测试题目和测试文本难度的因素，还探讨了阅读研究、阅读评价研究及阅读评价本质之间的关系。

第四章重点探讨了阅读的构念以及构念与测试规范之间的关系，阅读能力在语言水平量表中的表述，以及阅读构念与交际语言能力构念之间的关系。

第五章为“测试设计框架”，介绍了 Bachman & Palmer 框架，从阅读的环境特征、测试说明特征、输入特征、预期回答特征、输入与回答的关系等方面分析了该框架的特征及对评价的启示。

第六章介绍了真实世界中的测试，从阅读的环境特征、测试说明特征、输入特征、预期回答特征等方面介绍了四个评价案例，其中既有教育类的，也有非教育类的。

第七章为“阅读测试的方法”，主要介绍了完形填空和填空、选择题等客观测试方法，文本编辑（editing tests），以及一些综合性测试方法和非正式评价方式。

第八章为“阅读能力的发展”，介绍了一些测试框架、阅读量表和阅读测试系列，以期从多个角度说明阅读能力发展的特点。

第九章主要介绍了阅读过程的研究方法，如：内省法、即时回忆、语言失误分析等，介绍了阅读策略，并展望了计算机和网络测试的前景。



## 总体评价

对已有研究进行综述是本书的主要特色，全书通过对阅读研究的综述探索阅读结果和阅读过程、影响阅读的各种因素、阅读的本质及阅读能力的本质；通过对测试评价研究和实践的综述分析已有评价的优势与不足，探索测试与评价的未来发展方向。本书没有对概念作详细解释，如：什么是技能、策略和技巧？三者有什么区别？什么是评价？什么是测试？评价与测试是什么关系？真实性评价到底是怎样的评价？正如作者所说，他是在综述他人的研究，使用的也是所引用的研究中使用的概念。这就需要读者有一定的图式基础，对书中提及的概念和问题有自己的理解和思考。本书的目的不是解决阅读评价的各种问题，而是呈现各种可能，帮助评价者更清楚地认识所评价的阅读活动，探索各种评价的可能性，提高评价的效度、信度和公平性。

一般说来，评价至少包括形成性评价和终结性评价，形成性评价主要关注过程，而终结性评价主要指测试，重点关注结果。本书所探讨的阅读能力评价、技能评价、策略评价等主要用于阅读测试，作者所介绍的各种案例也多为

测试案例、测试量表或测试规范，可见本书主要关注的还是终结性评价。但是，作者也强调阅读评价还应关注阅读过程及学生阅读能力的发展。要促进阅读能力的发展，就需要开展形成性评价。关于策略评价、技能评价、阅读方式评价、影响阅读能力发展的因素的研究也都可以为阅读的形成性评价提供参考。形成性评价还需要诊断读者所处的发展阶段，以及他们是否达到了某阶段的能力要求。本书介绍的阅读过程、阅读结果等方面的研究可以为形成性评价和终结性评价的开展提供有益的参考。

随着评价理论的发展，真实性评价和任务型测试越来越为人们所青睐，因为它们代表了评价和测试发展的方向。真实性评价是“一种要求学生通过完成真实任务来展示对所学知识的掌握情况，以及对技能运用能力的评价方式”，真实性评价任务是“与成人、消费者或职业人士在工作中所面临的真实问题类似的任务”（Wiggins, 1993），这与本书中作者所倡导的根据目标语应用环境下的阅读任务设计阅读任务、撰写阅读问题是一致的。任务型测试是一种以任务为核心的测试方式。在任务型测试中，任务是分析的基本单位，测试题目的选择从任务出发，测试工具的编写以任务为基础，测试所要评定的是应试者在完成任务过程中的表现（Long & Norris, 2000）。本书倡导分析真实语言应用环境下人们的阅读情况、选择真实的文本、设计真实的任务，反映了测试发展的方向<sup>1</sup>。

---

<sup>1</sup> 本文引用人名和出处凡出自原著的，不再附带参考文献。

## 2 各章节主要内容与导读

---

### ●● 第一章 阅读的本质

要测试阅读，必须首先了解阅读的本质，本章有选择性地对阅读的本质进行了综述。

#### 一、过程与结果

阅读的过程是读者与文本互动的过程。阅读过程是动态的、变化的，同一读者在不同时间或以不同目的进行阅读时也会有不同感受。

对阅读过程的研究通常采用观察眼球运动、朗读中的错误分析及有声思维等方式。也有学者通过对阅读效果的评价研究阅读过程，但是，这种以结果为导向的研究方式有其局限性：一是结果的多变性，由于读者的经历、背景图式、知识和语言能力不同，对文本的理解就会存在差异；二是测量结果的方法不完善，如果评价涉及记忆，那么就很难区分最终评价的是理解还是记忆；此外，如果读者对评价方式不熟悉，评价方式也会成为影响理解的一个因素。

## 二、理解的层次

任何对文本的理解都表现为不同的层次,如:字面理解、推理理解及批判性理解等。一般认为推理理解比字面理解深刻,而较之这二者,人们更重视批判性理解。人们的这种价值判断造成了理解层次的划分。通常认为,读者首先学习字面理解,然后是推理理解,其后才是对文本的批判性理解和评价。但是,由于读者图式的差异,可能出现能够做出正确的推理但字面理解仍有困难的现象。因此,在设计阅读理解评价时,必须区分文本理解的层次,并关注读者图式的问题。

## 三、“能够阅读”的真正含义是什么?

有些读者能够理解单词却不一定理解句子,能理解句子却不一定理解文章的结构。Kintsch 和 van Dijk (1978) 认为前者与“微过程”相关,而后者与“宏过程”相关。微过程主要指局部的、词组层面的理解,宏过程指综合理解。研究者和理论家们把阅读能力分解为不同的阅读技能和微技能。从表面上看,这些技能的划分为测试任务的设计提供了理论支撑,且提供了诊断读者问题的可能。但是,作者认为,这些划分更多的是理论思考而非实验观察的结果;它们的定义模糊且相互交叉;有时很难辨别某一测试题测量的到底是哪种技能;对考试表现的分析不能体现阅读技能的可分离性和蕴涵量表,甚至不能体现阅读的难度或辨识能力的级别。尽管如此,用技能界定阅读的方法仍然很受欢迎、很有影响力,在探讨阅读本质的时候不能忽视。

目前的研究者越来越倾向于认为,阅读由解码(decoding)和理解(comprehension)两部分构成。解码是识别词汇,而理解则包括对句子进行语法分析、理解句子在语篇中的意义、构建语篇结构以及将这些理解与已有知识整合等。还有的研究者认为阅读由词汇识别技能、阅读速度和流畅性,以及问题解决和理解能力构成。

## 四、阅读时我们都在做什么?

有学者认为阅读过程包括先把印刷符号转变成语音符

号, 然后进行意义的识别, 也有学者认为读者直接处理意义, 不需要经过语音转换。尽管如此, 人们还是越来越多地认可阅读需要前期语音激活的理论, 认为语音识别是意义识别过程的一部分。阅读涉及大量的大脑活动, 有的活动是有意识的, 如: 决定只浏览报纸标题或读侦探小说时先看结局; 有的是自动的, 如: 阅读中读者很少意识到自己是如何将角色的面容、服饰、声音、行为地点及周围景观等视觉化的。

## 五、自下而上和自上而下

近二十年来, 研究者尤为关注读者采用的两种文本处理方式——自下而上和自上而下。自下而上的方法是一个序列模型。读者先识别图像刺激, 将其解码为声音, 识别单词, 然后才是理解, 后面的操作建立在前面操作的基础之上, 形成一个序列。

很多研究者更为强调读者已有知识(图式)在理解中的重要性, 认为阅读中读者需要激活其已有的相关图式, 构建阅读信息与已有图式的联系。自上而下的方式强调这些图式以及读者对阅读文本的作用。但是很多心理学家和心理语言学家怀疑图式理论对理解过程的解释作用, 最近的研究也因而更多地关注自下而上(或者信息驱动处理)对流畅阅读的作用。

其实, 不管是自下而上还是自上而下的方式都不能恰当地描述阅读过程。更为有效的模型是交互式模型。在这个模型中, 阅读过程的各个部分可以和任一其他部分互动。阅读过程的各组成部分不是序列性关系, 而是平行的(Grabe, 1991)。

这部分内容对评价的启示: 根据读者的模型和数据, 可以假设阅读能力不强的原因到底是自上而下的还是自下而上的。诊断性的测试者可以同时关注这两种可能性, 而不是只关注其中一种。可以诊断较差的阅读究竟是因为对自下而上的策略掌握得不好, 还是对背景知识了解得不够。

## 六、阅读与认知

根据 Goodman (1969, 1982) 对阅读过程的描述, 阅读过程其实是一种问题解决策略。Thorndike (1917) 则把阅读描述为一个推理过程。目前对批判性阅读技能发展的关注不仅利用阅读与思维领域的研究, 而且对二者几乎不加区分。在开发批判性阅读的测试之前, 必须弄清楚要在多大程度上评价阅读能力, 以及希望在多大程度上将阅读能力和其他认知能力区分开。如果要全面了解一个人不受其他认知变量干扰的阅读能力, 就应该努力将“阅读”和“推理”区分开。

## 七、母语阅读和二语阅读

阅读能力是否能在不同语言之间迁移? 很多二语教师都倾向于认为, 二语学习者阅读能力欠佳多是由于缺乏良好的母语阅读技能或习惯。Alderson (1984) 认为, 母语阅读能力要迁移到二语, 二语阅读者很可能必须先通过一个语言门槛。语言知识和阅读知识的重要性显然已被认可, 然而, 有证据表明, 在二语阅读中, 二语知识比母语阅读技能更重要。二语阅读能力差是由于母语阅读能力欠佳的观点没有得到研究文献的支持。二语阅读表现差很可能是由于语言知识不够所致, 想弥补的话, 可以多关注语言问题而非假定的阅读缺陷。

## 八、阅读作为社会文化活动

阅读通常是有目的的, 发生在一定的社会环境下, 社会环境也会影响读者对阅读本质的理解。最近的思想者倾向于将其表述为“能读会写”。现实世界充满了读写事件, 但是, 读写能力的使用、表现和评价会因文化的不同而不同, 对阅读的评价也因此会带有文化的偏见, 从阅读问题的设计到对测试结果的解读, 都被打上了文化的印记。

就阅读评价而言, 有学者认为应该开展阅读单项评价, 也有学者倡导听、说、写结合的综合性评价。但是, 如果读者不能回答问题, 应该注意区分是因为其表达能力不好还是理解能力不好。有人认为阅读评价本身规定了具体的测试项目, 是不恰当的; 还有人认为评价行为改变了读写活动的社

会文化属性，造成了阅读评价中的测试效应；任务的介入有违阅读的本质，剥夺了自然阅读的快乐。

## 九、对测试的启示

人们对阅读的认识将直接影响阅读的测试与评价。作者在回顾前人的研究，尤其是借鉴 Grabe (1991) 提出的阅读教学和阅读课程的指导原则的基础上，就阅读的测试与评价提出如下建议：

- 1) 可以用侧重内容的文本来评价阅读；
- 2) 应测试学生一系列的相关技能和策略；
- 3) 应鼓励学生阅读较长的文本，评价任务应能使學生体验阅读的快乐；
- 4) 应该用背景知识来促进阅读，而不是用背景知识的缺失来抑制阅读；
- 5) 测试结果应该有多种解释的可能；
- 6) 可以设计小组活动，讨论学生对文本的解读；
- 7) 评价过程不应忽视泛读；
- 8) 应该进一步探索识别技能的重要性，寻找测试这些技能的方法，并注意对阅读速度的测量；
- 9) 鉴于有些测试不可能不受外部因素的影响，对这些测试结果的解释应谨慎，如有可能，应参照其他评价方式解释测试结果；
- 10) 由于需要清晰地描述学生的阅读能力，对阅读或写作技能的综合性测试有时会无法实现；
- 11) 应探索相应的方法，对综合法、评价技能 / 策略、元认知知识及技能监控进行测试与评价；
- 12) 测试设计者应考虑，自己设计的测试在多大程度上立足于并反映了最近关于阅读过程（不仅是阅读结果）的研究成果。

## ●● 第二章 影响阅读本质的因素

关于二语阅读的研究越来越多，但是阅读评价者需要了解的是阅读本质的相关研究及发现。本章的主要目的是

让读者了解与测试和评价程序设计相关的研究成果。

## 一、读者因素

本部分主要介绍读者的知识、动机、策略、性别、个性等因素对阅读过程和阅读结果的影响。

### (一) 图式和背景知识

图式指的是代表读者知识的连锁的心理结构。读者处理文本时,会将文本中的新信息融入他们已有的图式,图式还会影响他们识别和存储信息的方式。对背景知识的作用进行解释的理论与图式理论稍有不同,但两种理论都强调读者知识的状态对阅读的过程、结果和回忆的影响。Carrell (1983a) 将形式图式和内容图式区分开来,前者主要指语言知识和语言学规范,包括文本的组织、某种体裁文章的主要特征等,后者可以划分为背景知识和主题知识。

### (二) 形式图式: 语言知识

语言知识亦即语言图式。语言图式对读者的阅读理解影响最大,内容图式次之(黄雪英,胡竹菊,2009);读者对文本信息的加工依赖于其语言知识。研究发现,读者需要认识文本中 95% 的词语才能做到充分的理解,才能根据上下文猜测词语(Laufer, 1989; Liu & Nation, 1985);读者做到熟悉文本中 97% 的词汇需要掌握大约 5,000 词汇。只掌握 2,000 最常用词汇的读者只能识别文本中 90% 的词汇(Hirsh & Nation, 1992)。

在二语和外语阅读研究中,总是假定读者必须首先获得语言知识后才能阅读。对二语阅读究竟是个语言问题还是阅读问题的研究提示,存在一个语言知识门槛(linguistic threshold),过不了这个门槛,读者就不能期望母语阅读能力迁移到二语阅读中。如果二语知识不够,就会影响成功的母语读者在二语阅读时的表现。二语的语言门槛因阅读任务的不同而不同,任务越难,门槛越高。

### (三) 体裁/文本类型知识

了解文本的组织方式、信息的表达方式、内容变化的标记方式等对阅读有促进作用。然而,对某体裁文本特征

的读者知识及其与阅读过程和结果关系的研究少之又少。可能很难将读者文本组织方式方面的知识与他们的元语言文本知识区分开。换句话说,读者可能知道文本的组织方式,但是他们没有意识到这种知识的存在,也不能清晰地将其表达出来。这就自然而然地引入了显性语言知识和隐性语言知识的讨论,在这儿,作者将二者分别称为元语言知识和元认知。

#### (四) 元语言知识和元认知

现代心理学家认为,元认知是“认知的认知”,是个体对认知过程的认识(Carrell, 1989)。研究发现,读者的阅读表现与其元认知水平有关。阅读能力差者不具备策略意识,经常意识不到该在何时以及如何应用自己的知识。元语言意识在学习阅读时发挥着作用。元语言知识与语言能力不同,教学生语言规则不太可能提高他们使用这门语言的能力。谈论文本语言的能力和理解文本的能力无关,后者在没有前者的情况下也是可能的。

#### (五) 内容图式

对内容图式的研究表明,读者需要了解与篇章内容相关的知识才能理解它。背景知识具有很强的效应。主题、结构、句法以及词汇相同的文本最容易激发读者已有的图式,使读者利用已有的知识进行阅读。

#### (六) 主旨/主题知识

如果读者对文本的主题一无所知,理解就会出现困难,对文本涉及的领域有所了解会有助于理解。图式理论的价值在于它试图解释新信息和旧信息的融合,但是很难解释新信息与不存在的信息是如何整合的。

#### (七) 世界知识

背景知识或世界知识对文本加工有一定的作用。所有的语言加工都需要世界知识。常识的激活是迅速的,并且是自动的,没有这种过程的话,理解速度会减慢,理解过程会变得艰难,也可能根本无法理解。

#### (八) 文化知识

文化知识对文本理解非常重要。内容和文化知识研究