

剑桥应用语言学丛书
CAMBRIDGE APPLIED LINGUISTICS

Language Program Evaluation

Theory and practice

语言课程评估

—— 理论与实践

Brian K. Lynch



上海外语教育出版社

剑桥应用语言学丛书

CAMBRIDGE APPLIED LINGUISTICS

Series Editors: Michael H. Long and Jack C. Richards

Language Program Evaluation

Theory and practice

语言课程评估

—— **理论与实践**

Brian K. Lynch



上海外语教育出版社

Language Program Evaluation

Theory and practice

Brian K. Lynch

The University of Melbourne



CAMBRIDGE
UNIVERSITY PRESS

图书在版编目 (CIP) 数据

语言课程评估: 理论与实践 / (英) 林奇 (Lynch, B. K.) 著.

—上海: 上海外语教育出版社, 2001

(剑桥应用语言学丛书)

书名原文: Language Program Evaluation: Theory and Practice

ISBN 7-81080-285-2

I. 语… II. 林… III. 英语—语言教学—教育评估—方法 IV. H319

中国版本图书馆CIP数据核字 (2001) 第068391 号

图字: 09-2001-161 号

出版发行: **上海外语教育出版社**

(上海外国语大学内) 邮编: 200083

电 话: 021-65425300 (总机), 65422031 (发行部)

电子邮箱: bookinfo@sflep.com.cn

网 址: <http://www.sflep.com.cn> <http://www.sflep.com>

责任编辑: 王彤福

印 刷: 常熟市印刷八厂
经 销: 新华书店上海发行所
开 本: 889×1194 1/32 印张 6.625 字数 284 千字
版 次: 2001 年 10 月第 1 版 2001 年 10 月第 1 次印刷
印 数: 3 500 册

书 号: ISBN 7-81080-285-2 / H · 109

定 价: 13.50 元

本版图书如有印装质量问题, 可向本社调换

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1996

First published 1996

Library of Congress Cataloging-in-Publication Data

Lynch, Brian K.

Language program evaluation : theory and practice / Brian K.
Lynch.

p. cm. – (Cambridge applied linguistics series)

Includes bibliographical references (p.) and index.

ISBN 0-521-48191-0. – ISBN 0-521-48438-3 (pbk.)

1. Language and languages – Study and teaching – Evaluation.

I. Title. II. Series.

P53.63.L96 1996

418'.007 – dc20

95-1719

CIP

A catalog record for this book is available from the British Library

ISBN 0-521-48191-0 Hardback

ISBN 0-521-48438-3 Paperback

Originally published by Cambridge University Press in 1996

This reprint edition is published with the permission of the Syndicate of the Press of the
University of Cambridge, Cambridge, England.

Licensed for distribution and sale in the People's Republic of China only and may not be
distributed and sold elsewhere.

本书由剑桥大学出版社授权上海外语教育出版社出版。

仅供在中华人民共和国境内销售。

出版前言

1999年5月至2000年6月间,上海外语教育出版社先后出版了从牛津大学出版社引进的“牛津应用语言学丛书”19种、“牛津语言学入门丛书”6种和“牛津应用语言学丛书续编”10种,受到了外语界师生的一致好评和欢迎。为了向我国的外语教学与研究人员提供更多的有关专著,帮助读者了解近年来国外应用语言学和外语教学研究的理论,促进我国外语教学研究水平的提高,上海外语教育出版社又精心挑选了剑桥大学出版社的应用语言学丛书10种,形成与“牛津应用语言学丛书”相辅相成的“剑桥应用语言学丛书”。相信这套丛书也同样能开阔我国学者的视野,通过借鉴国外研究成果,进一步总结我国自己的外语教学经验,形成具有中国特色的外语教学理论。

迄今为止,众多有关测试和评估语言学习者的著作已问世,但是探讨语言课程评估问题的专著不多。由墨尔本大学交际技能和ESL研究中心主任布赖恩·K·林奇博士(Brian K. Lynch)撰写的《语言课程评估:理论与实践》一书从宏观上探讨了语言课程的评估方法,填补了应用语言学论著在这方面的空白,向教师、语言研究人员和教育管理人员全面地介绍了课程评估的各种定性和定量分析的方法。作者将理论研究和教学实践有机地结合,从而提出了一个为语言教育课程度身定制的课程评估方法。该方法可以根据不同的情况进行调整,同时也可以随着教与学的变化与发展而逐步完善。

本书利用语言教学中的实例来说明课程评估的模式,并介绍了诸多数据收集和分析的方法。作者认为,课程评估应从只关注课程结束时考试成绩的纯定量研究转向包括调查课程实施过程在内的定性分析和自然研究方法。

全书共分九章。第一章根据评估要素适应模式概述课程评估

中的关键因素,以及课程评估与应用语言学研究之间的关系等。第二章综述从20世纪60年代至90年代中期采用定量或定性研究方法进行课程评估的研究项目,两种研究方法之间的不同点,以及当前的研究趋势。第三章从实例和自然研究法的角度阐述课程评估本身的效度。第四章着重讲述采用实例研究法的评估模式或研究方案。第五章主要介绍采用自然研究法的评估模式。第六章侧重讨论收集定量分析数据的各种方法。第七章讨论定性数据的收集方法,例如观察法、面试法等,并介绍了分析与解释定性数据的方法。第八章主要列举融实证研究法与自然研究法为一体的课程评估模式。第九章总结全书的主要内容。

本书的读者对象为英语教师、英语师资培训人员、英语语言课程负责人、教育管理或评估部门的人员,同时也可供外语专业研究生在学习应用语言学研究方法课程时使用。

for my mentor and friend, Russ Campbell,
and to the loving memory of my Dad

Series editors' preface

Program evaluation is important and difficult work in any field, and language education is no exception. The goal is sometimes to evaluate a program's effectiveness in absolute terms, sometimes to assess its quality against that of comparable programs, sometimes both. In ideal circumstances, evaluations receive cooperation from all parties and provide useful information to insiders on how their work can be improved, while offering accountability to outside stakeholders, such as host institutions, governments, and financial sponsors, as well as to students.

Circumstances are often less than ideal, however. Whether insiders or outsiders themselves, evaluators may be expected to employ recognized instruments and procedures, such as standardized proficiency tests and inferential statistics, for gathering and interpreting data. At the same time, they must also adapt to unique local conditions, where, for instance, objective measures yielding quantifiable data may be unwelcome, unusable, or unavailable. Worse, some stakeholders may have incompatible goals, conflicting interests in the outcome of an evaluation, and/or differing views about how it should be conducted. For example, an evaluation requiring full staff cooperation may be commissioned by a host institution, such as a university, with the aim of using the results to justify an already determined policy change with significant potential fall-out for program staff, including job losses. Training in conflict resolution may seem as useful as knowledge of applied linguistics in such cases, and the evaluator can easily end up taking sides or trying to play the role of mediator between warring parties.

Language Program Evaluation deals with these and many other methodological and policy issues, invoking both qualitative and quantitative, or positivistic and naturalistic, approaches in the process. The author, Dr. Brian K. Lynch, of The University of Melbourne, has had extensive field experience in both foreign and second language program evaluation and has made major contributions to theory and practice in this area, including his widely respected *context-adaptive model* for evaluation in applied linguistics. Dr. Lynch's book provides a coherent,

rigorous, and comprehensive approach to language program evaluation and also identifies contributions that program evaluation can make to the development of applied linguistics as a research field. It should appeal to a wide range of applied linguists and language educators and is a welcome addition to the Cambridge Applied Linguistics Series.

Michael H. Long
Jack C. Richards

Preface

This book had its beginnings in a motel room in Guadalajara, Mexico. It was there, planning the schedule for the soon-to-be-inaugurated Reading English for Science and Technology (REST) project at the University of Guadalajara, that I rediscovered a manuscript passed on to me by Russ Campbell. The manuscript was a copy of a paper delivered by Michael Long at a TESOL convention (later published in *TESOL Quarterly* [Long 1984]). Reading that paper provided the revelation that led to a preliminary design for evaluating the project, which became the focus of my dissertation, which, in turn, evolved over a six-year period into this book.

Drawing upon the experience of implementing and evaluating the REST project from June 1985 to July 1987 reported on in my dissertation, I constructed the initial outlines of a book designed to provide a thorough theoretical background for the evaluation of language education programs, as well as a guide for putting theory into practice. The earlier drafts of the book made use of insights and feedback gained from teaching a graduate seminar for applied linguistics students at UCLA and from my experience as Academic Director of the ESL Service Courses at UCLA, as well as from supervising graduate student research into program evaluation and serving as evaluator for the California Foreign Language Teacher Preparation Project.

The result of these experiences is a book that presents program evaluation against the historical backdrop of modern language teaching and the development of research in applied linguistics. In particular, the theoretical basis for language program evaluation is discussed in terms of the *quantitative-qualitative debate* in educational measurement, more recently referred to as the *paradigm dialog* (Guba 1990). Validity, the act of determining what counts as evidence for evaluation, is discussed from both the quantitative, positivistic perspective and the qualitative, naturalistic perspective. My goal in presenting this dialog and these, at times, conflicting perspectives, is to present the reader with the necessary range of theory and associated practical techniques that will allow for the most thorough response possible to any language program evaluation context.

The book uses the *context-adaptive model* for program evaluation (Lynch 1990) as an organizing principle for presenting the critical issues facing those who are about to embark upon the evaluation process. As with the context-adaptive model, the book was designed for those involved in language education. The concepts and methods advanced are illustrated with examples from my own experience, primarily ESL and EFL, as well as with vignettes from a variety of second language settings. It will be useful for language teachers who are looking for a systematic way to evaluate the instructional sequences of their classrooms, for language program administrators who need to decide on strategies for evaluating their schools, for language educators who have been asked to provide information on program processes or to make decisions about curricular change, for teachers of graduate level seminars on language program evaluation and research methods, and for a variety of applied linguists conducting research that focuses on, or makes use of, program evaluation data.

Whenever one manages to complete a project such as this, there are innumerable persons who have provided help, vision, encouragement, and advice along the way. Let me try to count and give recognition to at least some of the innumerable. Firstly, to all of those involved in the first two years of the REST project at the Universidad de Guadalajara (UdeG), especially my co-coordinator on the project, Margarita Matte, and the teachers on the project during the formative 1985–1987 years: Lyn Carr, Young Gee, Kathy Lunde, Suzanne McMeans, Carlos Oceguera, and Alfredo Urzúa. My special thanks go to the other three REST teachers from that period who have become lifelong friends: Elizabeth Borkowski, Juan Carlos Gallego, and Shira Smith. All of our students from the Chemical Sciences Faculty of the UdeG deserve recognition for their efforts in the program and as participants in the gathering of evaluation data. My thanks also go to the various officials at the UdeG and UCLA; in particular, Raúl Padilla, Margarita Sierra, Elwin Svenson, and Russ Campbell, whose unwavering support for the project was critical to its success.

I would also like to acknowledge Leigh Burstein, Russ Campbell, Evelyn Hatch, and Harold Levine for their helpful comments on the dissertation from which this book grew. I also wish to give special recognition to Harold Levine, who initiated me into the world of naturalistic research. That experience continues to be one of profound importance in my work. I was saddened to learn that Leigh Burstein passed away months before this book was due to go to press. Much of the knowledge of quantitative analysis that I present comes from my experience in his graduate seminar on quasi-experimental design.

When the book had taken sufficient form to begin the search for a

publisher, I received some invaluable criticism on the manuscript from Lyle Bachman, which allowed me to rethink the issues of content and audience in a productive fashion. At the same time, I received important feedback from my students and colleagues at UCLA who attended my seminar in program evaluation for applied linguists; in particular, Patsy Duff, Paul Gruba, Sally Jacoby, Andrea Kahn, Anne Lazaraton, Charlene Polio, and Jean Turner. The final time the seminar was offered, in the spring of 1992, I received helpful reactions to the content and detailed feedback on the manuscript from Bob Agajeenian, Denise Asher Babel, Linda Choi, Tom Griggs, Eric Miller, Micky Safadi, Kim Thomas, Jim Valentine, and Rosa Maria Victori-Blaya.

I would be remiss if I did not mention those who have provided the inspiration, friendship, and support during my years at UCLA, without which this book would never have survived the transition from experience to the printed page. A major expression of gratitude goes to my colleagues in the ESL Service Courses: Donna Brinton, Janet Goodwin, Christine Holten, and Sandy Wallace (also to Jean Turner, for her year as Academic Coordinator). To those in the Department of TESL and Applied Linguistics who were my teachers, I express my sincere thanks: Roger Andersen, Frances Butler, Russ Campbell, Marianne Celce-Murcia, Evelyn Hatch, John Schumann (and to the memory of John Povey). While they did not comment specifically on this manuscript, I owe a debt of gratitude to J. D. Brown, my original mentor who pointed me down the language-testing path and, again, to Frances Butler who took me further down the path with great insight and example. My thanks go also to Fred Davidson, for being such an enthusiastic and conscientious language tester and applied linguist. Most importantly, I want to thank Thom Hudson for his inspiring research and irreplaceable friendship.

In the final months leading up to publication of this book, I have received expert assistance and support from Jane Mairs, Mary Vaughn, and Mary Carson from Cambridge University Press. My thanks also go to Jack Richards, series co-editor, and to the anonymous reviewer of the manuscript for comments that improved the book in significant ways. At the same time, I have been welcomed to a new home in Melbourne, Australia by the students and staff of the Department of Applied Linguistics and Language Studies at The University of Melbourne. I give my special thanks to Gladys Cubberly and Vittoria Grossi for their help in that adjustment process, and to Tim McNamara for his encouragement, advice, and friendship.

During all of these phases I have received the ultimate support that only one's life partner can give. My love and thanks to Buni and to our son, Sam.

Finally, returning to this project's origins in that Guadalajara motel room, I thank Mike Long for inspiring the original concept, and for championing the manuscript all along the way. I will end this lengthy list of thank you's with an acknowledgment of the profound example that has been provided for me and for the field of applied linguistics by Russ Campbell. Russ represents the best qualities of a teacher, administrator, researcher, mentor, and friend. He is the heart of this book.

Brian K. Lynch

Contents

Series editors' preface	ix
Preface	xi
1 Introduction	1
Definitions: Applied linguistics, evaluation, program	1
Critical issues	2
Program evaluation and applied linguistics research	9
2 Historical background	12
The paradigm dialog	12
Evaluation of language education programs in the 1960s and 1970s	22
More recent developments in the evaluation of language education programs	32
Summary	39
3 Validity	41
Validity from the positivistic perspective	41
Validity from the naturalistic perspective	53
Conclusions: Validity from the two perspectives	66
4 Positivistic designs	70
True experimental designs	71
Quasi-experimental designs	73
5 Naturalistic designs	80
The responsive model	80
The illumination model	82
Goal-free evaluation	84
The judicial model	85
The connoisseurship model	86
Other metaphors for naturalistic evaluation	88

6	Quantitative data gathering and analysis	92
	Data gathering	92
	Data analysis	94
	Conclusion	105
7	Qualitative data gathering and analysis	107
	Overview	107
	Data gathering	108
	Data analysis	139
8	Combining positivistic and naturalistic program evaluation	155
	Compatibilist versus incompatibilist perspectives	155
	Mixed strategies	156
	Multiple strategies	158
	Mixed designs	159
	Mixed designs and strategies over time	165
9	Conclusions	167
	CAM step 1 (audience and goals): Determine the purpose of the evaluation	167
	CAM steps 2 and 3 (context inventory and preliminary thematic framework): Determine what is being evaluated	170
	CAM steps 4 and 5 (evaluation design and data collection): Select a design and collect the data	171
	CAM step 6 (data analysis): Analyze and interpret your findings	173
	CAM step 7 (evaluation report): Communicating the evaluation findings	174
	The role of program evaluation in applied linguistics research	175
	References	178
	Author index	188
	Subject index	191

1 Introduction

It is probably safe to assume that the concept of program evaluation is not completely foreign to most applied linguists, even to those working outside the language education domain. Certainly the words *program* and *evaluation* conjure up reasonably clear mental images, and the notion that a program might need to be evaluated does not seem illogical to most. Language education programs abound internationally, and the majority of applied linguists have most likely, at some stage in their career, been involved in these programs as teachers, administrators, students, researchers, or some combination of these roles. Many, if not most, have been involved in some sort of effort to evaluate a language program. This evaluation may have taken the form of asking students to rate their language course and teacher using a questionnaire, giving achievement tests at the beginning and end of a period of instruction, or having a language teaching expert from another institution visit the program and prepare a report on its strengths and weaknesses. Program evaluation, then, can be seen as relevant to the experience of a wide range of applied linguists, and will be of particular interest to language educators.

Definitions: Applied linguistics, evaluation, program

In order to proceed with a detailed examination of the theory and practice of program evaluation within the broad context of applied linguistics, however, more precise definitions of certain terms are in order. I will focus on three key terms here; others will be presented in Chapter 2. To begin with, *applied linguistics* (AL), as a term defining an emerging academic discipline, has been the subject of recent discussions (Andersen et al. 1990; Pennycook 1990; van Lier et al. 1991). For the purposes of this book, AL will refer to research and practice concerned with the application of knowledge and methods from a variety of disciplines (e.g., anthropology, sociology, linguistics, psychology, and education) to the range of issues concerning the development and use of language.