Talking About Language Assessment: The LAQ Interviews

Edited by
Antony John Kunnan

TALKING ABOUT LANGUAGE ASSESSMENT

The LAQ Interviews

Edited by Antony John Kunnan

Routledge
Taylor & Francis Group

The right of Antony John Kunnan to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

TALKING ABOUT LANGUAGE ASSESSMENT

Talking About Language Assessment is the first book to take an interview-based approach to the history of language assessment. This collection consists of interviews originally featured in *Language Assessment Quarterly* with 12 renowned language assessment experts: Alan Davies, Bernard Spolsky, Merrill Swain, John Carroll, John Trim, Lyle Bachman, Charles Stansfield, Gui Shichun, Kenji Ohtomo, J. Charles Alderson, Elana Shohamy, and Liz Hamp-Lyons. Collectively, this group of academics makes up some of the most highly cited and distinguished professors and researchers in the world from the area of language assessment and the related fields of language acquisition and language policy. Contextualized with commentary by the editor, the chapters in this text cover many of the essential areas of language testing and assessment including assessment design, validation argument, epistemological issues in research, and language assessment policy. This text is a valuable compilation on the history of language assessment, and an ideal resource for researchers and students in the field of language testing and assessment.

Antony John Kunnan is a professor of English at the Nanyang Technological University, Singapore. He is the editor of *The Companion to Language Assessment* (2014), the founding editor of the journal *Language Assessment Quarterly* (2003–13), as well as past president of the International Language Testing Association.

New Perspectives on Language Assessment

Series Editors: Antony John Kunnan, Nanyang Technological University, Singapore, and James E. Purpura, Teachers College, Columbia University, USA

Headed by two of its leading scholars, this exciting new series captures the burgeoning field of language assessment by offering comprehensive and state-of-the-art coverage of its contemporary questions, pressing issues, and technical advances. It is the only active series of its kind on the market, and includes volumes on basic and advanced topics in language assessment, public policy and language assessment, and the interfaces of language assessment with other disciplines in applied linguistics. Each text presents key theoretical approaches and research findings, along with concrete practical implications and suggestions for readers conducting their own research or developmental studies. A companion website to the series houses such features and tools as book chapter outlines and summaries, sample testing materials, test performance data, activity sheets, quizzes, annotated bibliographies, web links to relevant institutions, and much more.

The Diagnosis of Reading in a Second or Foreign Language
J. Charles Alderson, Eeva-Leena Haapakangas, Ari Huhta, Lea Nieminen, and Riikka Ullakonoja

Talking About Language Assessment: The LAQ Interviews
Edited by Antony John Kunnan

SERIES EDITOR PREFACE

Language assessment or testing started in the late 19th century with assessments in foreign languages (French, German, and Spanish mainly) in the U.S. and English as a second language at Oxford and Cambridge Universities in the U.K. and language assessments in France (*le baccalauréat*) and Germany (the *Abitur*). But most scholars consider the birth of modern language testing as a field of study to be the year 1961, with the publications of Robert Lado's book *Language Testing* and John Carroll's chapter "Fundamental Considerations in Language Testing," and the earliest efforts to organize one of the world's influential tests—what is known today as the *Test of English as a Foreign Language*, or *iBT (internet-Based TOEFL)*. In the last 50 years, the field has developed into an interdisciplinary enterprise, in part influenced by disciplines such as applied linguistics, language teaching, educational and psychological measurement, and, more recently, sociology and ethics, and in part by carving out its unique identity through the development of its own theories and practices.

This series will capture this burgeoning field by offering a cogent and comprehensive state-of-the-art coverage in the following areas:

1. The use of quantitative and qualitative research methods for conducting research into various aspects of assessment (for example, validation, absence of bias, accommodations, administration);
2. The development and use of language assessments to understand and promote learning and performance in diverse contexts including traditional classrooms (for example, diagnostic feedback, placement);
3. The new areas of language assessment that are becoming critical to the field such as pragmatics, translation, intercultural communication (for example, in contexts such as assessing aviation specialists, court interpreters, and translators);

4. The modern use of technology in terms of computer-delivered and computer-adaptive tests, and speech recognition and automated scoring of writing;
5. The development of language assessments for immigration, citizenship, and asylum contexts.

The primary purpose of this series is to provide target audiences with theory, research, and practice-informed academic texts for addressing some of the contemporary questions, issues, and concerns in the field that involve the use of assessment. The scope of the series is designed intentionally to be broad enough to include books that discuss research methods for carrying out language assessment research as well as newer uses of language assessments and technology in the field of applied linguistics.

The secondary purpose of this series is to serve a very broad international language assessment community. The series will also feature languages other than English and the English-speaking world (through examples and practices from other parts of the world) as we believe insights from international contexts and multiple languages can potentially inform the field.

This series will be of interest to students of language assessments at undergraduate, graduate, and doctoral programs; faculty in applied linguistics; as well as officers in education, government, and military departments involved in language assessment. Researchers in other fields of applied linguistics (e.g., discourse analysis, L1 literacy, SLA, second and foreign language pedagogy, and communications) and psychology, education, measurement, assessment, and research methods will also find the series valuable.

Book Preface

In the current book, *Talking About Language Assessment: The LAQ Interviews*, Antony John Kunnan, founder and former editor of *Language Assessment Quarterly*, has assembled the 12 interviews from *LAQ* into one volume. The interviews were conducted with J. Charles Alderson, Lyle F. Bachman, John B. Carroll, Alan Davies, Liz Hamp-Lyons, Kenji Ohtomo, Gui Shichuan, Elana Shohamy, Bernard Spolsky, Charles Stansfield, Merrill Swain, and John Trim. This volume begins with an interesting introduction by Antony John Kunnan and Nick Saville, the *LAQ* associate editor charged with commissioning the interviews at the time, in which they discuss, in question-and-answer form, the logistics of selecting interviewees and producing an interview. The interviews are then presented in their original form, with some minor updates and modifications.

This volume provides readers with the rare opportunity of getting to know some of the most distinguished thinkers and prolific researchers in the field of language assessment between 2004 and 2014, when the interviews were first published in *LAQ*. The volume allows readers to see each interviewee's language assessment story as the interviewees reveal what they were thinking at the time and what

led them to investigate the intellectual questions addressed in their work. Finally, through the interviews, we hear oral accounts about the design of language assessments, language assessment in the developing world, the creation of the Common European Framework of Reference, validation theory and practice, and language assessment policy. This book is a uniquely personal contribution to the history of language assessment.

James E. Purpura
Siena, Italy

INTRODUCTION
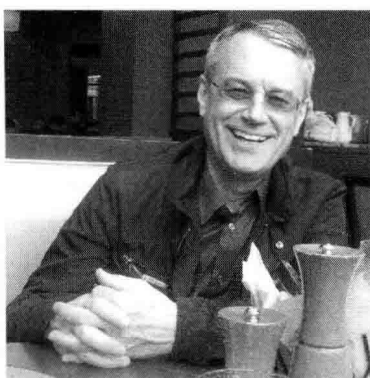

FIGURE 1 Antony John Kunnan

FIGURE 2 Nick Saville

In this volume we have a wonderful collection of 12 interviews with arguably the most important experts in language assessment at the start of the 21st century. The experts have university affiliations in the U.S., U.K., Canada, China, Japan, and Israel but are internationally known for their distinguished scholarship and service to the field of language assessment over the last 50 years. Nick Saville, as associate editor of *Language Assessment Quarterly*, was the commissioning editor for this series of interviews. He and Antony John Kunnan, editor of *LAQ*, talked about the many aspects of the collection in the manner of a conversation. We present this in a question-and-answer format.

Q: What was the purpose of the interviews?
A: The interviews were part of the "new directions" being taken by the editors to expand the coverage of genres in the newly founded *LAQ*. The aim was to

identify key figures in the field—and through an interview format to produce a dialogue with them that would reflect their background and experiences with a focus on language testing. This would reveal what had led them into the field and how their formative experiences fashioned their views and, ultimately, their academic work and contributions to the field.

The narrative/discursive style would add new perspectives on well-known events, people, and trends in the field—thus contributing depth or nuance to theoretical and practical concerns. You will recall that this happened before the current era of YouTube and "always on smartphones"—although videos were already available of well-known figures talking about their work. The editors wanted to capture the voices of the interviewees in print—so the aim was to capture the style and content which would be recognized as typical of the speaker by anyone who knew them well.

Q: Was the focus on the accomplishments of the interviewee?
A: The focus was not so much on accomplishments as experiences and perspectives—although the main achievements were usually covered to some extent.

Q: Was an attempt made to record oral histories of the time period of the interviewees and their work?
A: The concepts of "voice" and of "being a witness" to events of historical interest are increasingly used in putting together and building up the historical record. Oral histories and narratives add nuances and insights not captured in other ways. They complement other methods of recording the past.

Q: How were the interviewees and interviewers selected?
A: The interviewees were selected using a number of criteria that were relatively loosely applied—with opportunity and practicality intervening from time to time. We were concerned to get a range of voices from different parts of the world, representing varied perspectives on the topics of interest, and of course with a balance of gender. It was inevitable that the interviewees would be senior members of the assessment or applied linguistics community and usually with a background in several academic and/or geographical contexts. In some cases they were chosen to coincide with an important event, such as a significant birthday or the announcement of an academic award.

We were also concerned that there should be a rapport between the interviewee and interviewer so that the interview itself when transcribed and edited would represent a permanent record of a unique event (in terms of time and place and between two or more people). Knowing the interviewee and the possible range of topics was an important consideration in selecting interviewers, as well as geographic proximity or attendance at the same event somewhere in the world.

Q: What format was used for the interviews? Question-answer or conversation?

A: The aim was to create a naturally flowing dialogue based on a "shooting script" of areas to be covered that had, broadly speaking, been agreed on in advance. Inevitably Q&A was part of this, but often the dialogue also involved the interviewer contributing views and opinions for the interviewee to react to or comment on.

Q: How were the interviews conducted or recorded? Face-to-face, a combination of face-to-face and e-mail, e-mail only?

A: The preference was to conduct the interviews face-to-face and to record them with a handheld device of some kind. This was not always possible due to practical constraints—the "best made plans etc." . . . Some interviews were conducted mainly online. Inevitably there was some preparation and follow-up to the interviews using remote communications, mainly e-mail.

Q: How narrowly or broadly were the interviews transcribed? Were the interviewees given opportunities to review the final copy?

A: The transcripts were broad representations of what was said—so not transcribed for conversation analysis. However, the aim was to capture the style of the conversation and not to create a text based on written language criteria. Again we wanted the reader to hear the voices of the participants when reading the interview in the journal.

   The interviewees were given the chance to comment, and in some cases clarification was sought on what was said—for example, when reference was made to people, events, or publications. In some cases, notes were added to make these references clear to a wider audience. In some cases, the interview was too long to include in its entirety and so had to be edited down. This was usually done by consent involving the interviewer, interviewee, and an editor. Sensitivity was required to ensure this kind of editing did not interfere with the flow or change the character of the interview by skewing the coverage of topics or cutting out key parts of a narrative.

Q: As a collection of 12 interviews, what does the book bring together?

A: The 12 interviews constitute a unique record, collected over a decade at the start of a new century. They provide reflections on the state of the art in our field as perceived by some of the most influential figures and bring into evidence personal views and opinions which otherwise would not be in the public domain.

We would like to thank all the interviewees and the interviewers for their dedication to this project; for planning and taking part and conducting the conversations; for editing, transcribing, and reediting the conversations; and for proofreading and making final corrections. We hope we have succeeded in our endeavor.

<div align="right">

Antony John Kunnan and Nick Saville
San Gabriel and Cambridge, July 2014

</div>

CONTENTS

The Interviews

1

JOHN B. CARROLL (1916–2003)

The Story Behind the Modern Language Aptitude Test

Interview by

Charles W. Stansfield

SECOND LANGUAGE TESTING, INC.

Daniel J. Reed

SECOND LANGUAGE TESTING, INC. AND INDIANA UNIVERSITY

PHOTO 1.1 John B. Carroll

The following is an interview with Dr. John Bissell Carroll, who was considered by many to be the premier psychologist in the 20th century in terms of contributions to educational linguistics. In retrospect, this occasion has very special significance, as it was one of the last interviews that Dr. Carroll granted near the end of his illustrious career.[1]

Dr. Carroll was lead developer of the Modern Language Aptitude Test (MLAT). The MLAT has been used for selection, placement, and guidance by schools and U.S. government agencies for more than 40 years, and Dr. Carroll's associated research and writings on language aptitude are widely considered the most comprehensive and authoritative treatment of the subject even today. Yet, not all of the story behind the MLAT has been previously recorded. One of the main purposes

of this interview was to augment the documentation of this important part of language testing history.

John B. (Jack) Carroll was born in 1916. During his long and distinguished career, he made many contributions to the study of Intelligence, verbal abilities, educational psychology, factor analysis, educational measurement, and language testing. In August 2002, just a few months before this conversation, Carroll traveled to Chicago to attend the American Psychological Association's 110th Annual Convention where he received the American Psychological Foundation's "Gold Medal Award for Life Achievement in the Science of Psychology" (Carroll, 2002). At the time of his death, he lived in Fairbanks, Alaska with his daughter, who is a Russian–English interpreter, and her husband, who is a professor of ecology at the University of Alaska. The interviewers, Stansfield and Reed, are affiliated with Second Language Testing, Inc., which republished the MLAT in January 2000 and the elementary school version, the MLAT-E, in fall 2002.

The MLAT was originally published in 1959, and although it has remained in continuous use ever since, the perceived appropriateness of the test faded for a time, at least in fields such as second language research and pedagogy, as more "communicative" approaches to second and foreign language learning were developed. Remarkably, however, the test has proven to be relevant even in the most current teaching and research contexts. For example, validity coefficients based on data collected in the communicatively oriented classrooms of the U.S. Foreign Service Institute are approximately at the same levels as reported in the original version of the test manual, largely in the .40 to .60 range (Ehrman, 1998). In fact, there has been a resurgence of interest in the aptitude construct and related research, some of which is reported in a recent book edited by Robinson (2002). Readers might also wish to compare these more recent treatments of aptitude to Carroll's earlier vision of how aptitudes, general intelligence, motivation, quality of instruction and other variables interact in school learning (Carroll, 1963). The MLAT has also been employed extensively in the study of learning disabilities that appear to affect language learning, and in the exploration of the possible existence of a "foreign language learning disability."[2] Although conceptualizations of language aptitude are evolving, much of the new research builds on, or supplements, the original notion, rather than refuting it. Skeehan summarizes his own conclusions, which presumably are shared by other researchers, by saying "traditional views of aptitude are still relevant to applied linguistics and SLA . . . [and] post-MLAT research into aptitude . . . has added to our understanding in significant ways" (as cited in Robinson, 2002, p. 92).

To learn as much as possible about the history of the MLAT, the interviewers arranged a 45-min telephone conference call with Carroll on October 30, 2002. The recorded call was transcribed, and a version of the transcription with footnotes and references, edited by Carroll and the interviewers, is presented here. The interviewers and the interviewee are identified by their initials, prior to each turn in the conversation. Notes and references provide additional background information.