

A Critical Investigation of Test Use: Language Testing, the Social Dimension

Reviewed by: [Jamie L. Schissel](#)

Schissel, J.L. (2010). Review of the book: Language testing: The social dimension. by T. MacNamara and C. Roever. *Linguistics and Education*, 21(2), 126-128. doi: 10.1016/j.linged.2010.05.004

Made available courtesy of Elsevier: <http://www.dx.doi.org/10.1016/j.linged.2010.05.004>

***© Elsevier. Reprinted with permission. No further reproduction is authorized without written permission from Elsevier. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. ***

This is the author's version of a work that was accepted for publication in *Linguistics and Education*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Linguistics and Education*, Volume 21, Issue 2, (2010) DOI: 10.1016/j.linged.2010.05.004

Abstract:

This article is a review of the book "Language Testing: The Social Dimension" by Tim MacNamara and Carsten Roever.

Keywords: Book Review | Language | Testing | Social Dimensions

Article:

Recipient of the Sage/International Language Testers' 2009 book award for language testing, Tim McNamara and Carsten Roever's *Language testing: The social dimension* discusses the increasingly relevant implications of language tests for individuals and society. As the title suggests, this book takes a more critical position on the role language tests play in society as opposed to a more traditional testing perspective. This approach focuses on the impact of tests, which builds on the work of Shohamy (2001) who argued for language testers to develop more strategies to investigate the consequences of test use. The authors have produced a book well grounded in the field of language testing that informs future directions for research and studies for any persons interested more broadly in assessment, or specifically in language testing. Their analyses of theoretical and empirical research in each chapter as well as their metacommentary about how this research relates to the field of language testing provides the reader examples that clarify the authors' positions. The historical background about other researchers' approaches is particularly useful for engaging meaningfully with the content of the book. Possible readers include researchers, graduate students, test developers and school administrators. In relation to

issues of testing and education, the authors' analyses offer information about test use in multiple settings (e.g., the workplace, for citizenship, in schools) that are directly and indirectly connected to educational contexts. The book has eight chapters and begins with an introduction to the book before delving into specific issues in the next six chapters. The final chapter synthesized the major arguments of the previous chapters and calls for changes in research approaches and the implications of implementing these changes.

Chapter 1 serves as a preview of the contents of each following chapter while also introducing some of the major questions and arguments posed by the authors. This chapter provides a brief overview progression of research in language testing, which is elaborated on further in each of the following chapters. The authors begin their detailed discussion on the social dimension of language testing in chapter 2 by reviewing how validity has been conceptualized from the 1950s to present day, especially in relation to test use. Samuel Messick's seminal work on validity is foregrounded through an examination of his predecessors and language testing researchers that have developed his ideas further (e.g., Bachman, Kane, Mislevy). Additionally, the authors point to limitations in the current research on validity, namely that there is no clear way to account for the influence of the broader social context in test use. This chapter, thus, looks back at previous work in order to determine which steps are feasible for moving forward to create an account of test use.

Chapter 3 is the first step in pushing research to go further by asking how testable the social dimensions of language proficiency are. This builds on chapter 2's discussions on construct validity by examining how proficiency is defined and measured. McNamara and Roever introduce strengths and weaknesses of testing face-to-face interactions, pragmatics and discourse completion, which are forms of assessment that are more related to the social use of language. They suggest that the complex, co-occurring events that characterize language-in-use may not be best accounted for in a controlled testing setting. Although the question posed at the onset of the chapter was not given a definitive answer, the arguments presented place a more critical lens on this issue as a means to proceed in a new direction.

After focusing on the social dimension of proficiency, the authors shift in chapter 4 to a psychometric perspective by providing a detailed explanation of approaches used to investigate test fairness taken by researchers in measurement. The bulk of the chapter discusses the analysis of bias through differentiated item functioning (DIF), a method that determines which test questions are answered differently by two different groups (e.g., in terms of sex, a question that males score significantly higher/lower on than females). Their description of DIF, including the history, strengths and limitations is approachable for readers with limited background in measurement or statistics. Taken on its own, this chapter is a comprehensive introduction on many psychometric methods that could be used to introduce students to these topics while also providing tools to read this research critically. Within the larger discussions of this book, this chapter clearly shows the role that psychometrics can and cannot play in understanding the consequences of test use.

Chapter 5 broadens its study of issues of fairness from the testing instrument to also include professionals in the language testing field. The authors examine various testing or testing related organizations' codes of ethics and fairness reviews. These documents detailing ethical practices for testing professionals represent an additional, but limited, approach taken by language testing researchers to prevent or account for negative consequences test use that recognize the political nature of tests. To clarify what they see as negative consequences, the authors apply the concepts from the previous chapters in the next two chapters.

Chapters 6 and 7 use a social theory lens to explore the various unintended consequences of test use. Investigating the interaction of identity and tests, the assessments featured in chapter 6 closely deal with issues surrounding the denial of rights especially in relation to asylum and immigration. Additionally, many of the immigration issues in chapter 6 have been expanded into a 2009 special issue of *Language Assessment Quarterly* titled: "Language Assessment Immigration, Citizenship, and Asylum." In the cases presented here, the authors also discuss how issues of identity and tests are linked to policies. In Canada, for example, the country's official bilingual policy has implemented the French and English language proficiency test of civil servants. Persons who successfully pass both tests represent an identity of a new Canadian that is promoted by the official language policy. The tests in this and many instances serve to reinforce the ideas that the policy promotes. Thus, the authors argue that ultimately "[t]ests create the identities they measure" (p. 196).

Focusing on language assessments at schools, chapter 7 foregrounds the politicization of assessments. Highlighting the Japanese high school entrance exam for English, the Common European Framework of Reference and *No Child Left Behind* (NCLB) in the U.S., among others, the authors argue the issues surrounding tests in these settings reach beyond their intended purpose. They offer a critique of the standards movement and contend that many standards of learning are derived from policy goals rather than developed from empirical research. The problem then arises that assessments must measure how students achieve these standards, and thus the tests constructs are based on policy, not research. Additionally, they argue that the impact of language assessments in relation to schools serves to meet broader cultural, political and social goals. In the case of NCLB, the impact on linguistic minority students is discussed largely as negative but with some potential for positive consequences. After the publication of this book, there has been continued research documenting the impact of NCLB assessments (see Menken, 2008) that supports their assertions that the intended and unintended consequences of this policy extend beyond schools.

The final chapter calls for more diversity in research, including Vygotskian influenced dynamic assessment (see Poehner, 2008) or assessments that reject standard language varieties in favor of, for example, English as a lingua franca. This broadening of the field of language testing also entails more cross-disciplinary collaboration in contrast to previous emphases on strict psychometric training.

McNamara and Roever's examination of the social dimension of language testing is comprehensive and introduces the reader to multiple issues that have become more relevant since the publication of this book, especially for those in the field of education where accountability systems that are reliant on tests have gained prominence. The questions posed and arguments raised are challenging and reflect a change of tenor of the field. Even if the reader does not agree with the claims made by the authors, the information provided in this book serve as a meaningful foundation for an insightful discussion and possible inspiration for moving in new directions in this field.

References

Language assessment for immigration, citizenship, and asylum. (2009). *Language Assessment Quarterly*, 6(1), (special issue).

Menken, K. (2008). *English learners left behind: Standardized testing as language policy*. Clevedon: Multilingual Matters.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskyan approach to understanding and promoting L2 development*. New York: Springer.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow, Essex: Pearson Education.