



PROJECT MUSE®

Language Testing Reconsidered

Fox, Janna, Wesche, Mari, Bayliss, Doreen, Cheng, Liying

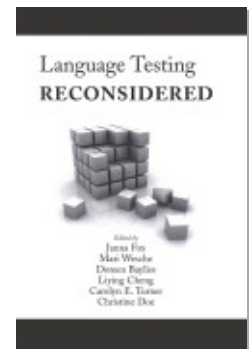
Published by University of Ottawa Press

Fox, Janna & Wesche, Mari & Bayliss, Doreen & Cheng, Liying.

Language Testing Reconsidered.

Ottawa: University of Ottawa Press, 2007.

Project MUSE., <https://muse.jhu.edu/>.



➔ For additional information about this book

<https://muse.jhu.edu/book/4459>

7

LANGUAGE TESTING:
A QUESTION OF CONTEXT**Tim McNamara***University of Melbourne**Abstract*

Arguably the greatest challenge facing language testing is the issue of the context in which language testing is carried out, both at the micro and the macro level. In standard works on language testing, context has been theorized in terms of the demands it makes on individual cognitive attributes, but this has distorted the picture of the social context available to us. The social context theorized in its own terms has featured rather weakly in discussions of language tests. At the micro level, insights from Conversation Analysis have challenged the individualistic focus of current thinking about assessment. At the macro level, the awareness of social context has led to the development of critical language testing and an understanding that language testing constructs, implicitly or explicitly, may be sociolinguistic in origin, in the context of the use of language tests as markers of identity in settings marked by intergroup conflict. The paper argues that all language tests can be seen as tests of identity in the light of the theory of subjectivity proposed by Foucault, using as an example the Occupational English Test, a test of English as a second language for immigrant health professionals. The paper concludes with an argument for a deeper engagement with contemporary social theory in language testing research.

In this retrospective paper I advocate a renewed discussion of the social context of language testing. This involves investigating the claim that all language tests are tests of identity. Such a claim is based on the reconceptualization of identity as subjectivity within recent social theory. In the discussion, I will refer to the Occupational English Test (OET) (McNamara, 1996), an Australian test of second language proficiency for immigrant health professionals.

The themes of the paper reflect the development of interests I had at the time I first attended the Language Testing Research Colloquium (LTRC), in San Francisco in 1990. At that LTRC I gave a paper reporting some of the findings of my PhD research on the validation of the OET, which I was at that time in the throes of completing. This was in fact my second PhD topic; the first, on which I had worked on and off for four years, had been a study of language and identity, but midway through I abandoned it and changed my focus to language testing, largely for pragmatic, job-related reasons. Recently, I have begun to teach again in the area of language and identity, and I have

found much in that field to stimulate, challenge, and deepen my understanding of language testing as a social practice. I would like to present some of this recent thinking in this paper.

How can language tests be seen as tests of identity? How does this apply to proficiency tests such as the OET? The function of certain language tests as techniques for identification is of course recognized in the literature on language testing; for example, in the frequent references to the Biblical shibboleth test, in which a single pronunciation feature was used to detect defeated enemy soldiers trying to “pass” among the victors (Spolsky, 1995). This function is currently of great interest in the recent critical discussion of the growing use of language tests within the asylum process (Eades, Fraser, Siegel, McNamara, and Baker, 2003; Reath, 2004; Language and National Origin Group, 2004; Eades, 2005). Determining which side of a political border a person is from, sometimes on the basis of a language test, may result in the individual being granted or denied refugee status under international law. Asylum seekers are interviewed by an immigration officer either through an interpreter or through the medium of a lingua franca, and a tape recording of the interview is subjected to language analysis, on the basis of which a claim is made as to the sociolinguistic and hence national identity of the speaker (Eades, 2005). The goal of identity tests such as this, and aspects of the procedures involved, seem obviously different from the goals and procedures of language proficiency testing. However, I want to show that by rethinking the *social context* of language tests, the distinction between tests of proficiency and tests for identification disappears. All language tests are about identification.

Before developing this point further, I want to briefly consider the way in which the social context of the test is conceptualized within the standard works on proficiency testing, for example in Bachman’s (1990) *Fundamental Considerations in Language Testing*. We need to distinguish context at two levels: the micro, that is, the immediate context, such as the context of performance in face-to-face speaking tests, and the macro, or wider social context. The discussion of context in Bachman (1990) focuses on the micro level. The emphasis in this discussion, a reflection of the broader traditions both of linguistics and educational measurement in which Bachman’s work is firmly located, is cognitive and psychological, even if the definition of the context with which language proficiency must engage is described in part in terms made available from the ethnographic work of Dell Hymes. Context is addressed most clearly in the notion of the target language use situation, which is conceived from the perspective of the language user; this reflects a wider needs-based emphasis in communicative language teaching. The target language use situation is defined in terms of its demands on the language user; this allows the test method facets — the test content and the testing procedures — to represent the interface between the dimensions of communicative language ability and the demands

of the target language use situation. In order for this set-up to work, the demands of the context are necessarily understood as cognitive demands. There is thus a feedback loop between context and ability: the target language use situation is conceptualized in terms of components of communicative language ability, which in turn is understood as ability to handle the target language use situation. The situation or context is projected onto the learner as a demand for a relevant set of cognitive abilities; in turn, these cognitive abilities are read onto the context.

What we do not have here is a theory of the social context in its own right, as it were, that is, a theory that is not primarily concerned with the cognitive demands of the setting on the candidate. In fact, a fully developed social theory of the face-to-face context does exist in the form of Conversation Analysis (CA). While this has been the basis for an increasing volume of research in language testing (Young and He, 1998; Brown, 2003, 2005; Brown and McNamara, 2004), the theoretical dilemmas involved in the use of CA in this context are far from resolved. Beyond this, the wider social context in which the immediate target use situation is located has been less well theorized. While the social implications of test use have been considered in the language testing literature at least since Spolsky (1981) and are discussed, albeit rather briefly, in Bachman (1990), discussion is restricted to the intended and unintended consequences of the test, not to the wider social meaning of the test in its context. There is a gap, in other words, in our theorizing of the social context in which language tests find their place.

Contemporary social theory offers a rich array of conceptual frameworks in terms of which one could potentially conceptualize the social context and role of language tests. One particularly relevant one, which additionally addresses the question of examinations, is in the work of Michel Foucault. Foucault addresses the way in which social identities, what he calls subject positions, become available as a function of dominant discourses ("big D" discourses, to use Pennycook's 1994 term) at any given historical moment. In his book *Discipline and Punish* (Foucault, 1977 [1975]), he sees the discourses of modernity which emerged in the late 18th and early 19th centuries as involving the instilling of a sense of self through techniques of surveillance, primary among which is the examination. In Foucault's brilliant analysis of the relationship between tests and power, he shows us how tests can be experienced as exercises in subjection to power, a process that is productive of individuality or, as he terms it, subjectivity:

The examination as the fixing, at once ritual and 'scientific', of individual differences, as the pinning down of each individual in his own particularity ... clearly indicates the appearance of a new modality of power in which each individual receives as his status his own individuality, and in which

he is linked by his status to the features, the measurements, the gaps, the 'marks' that characterize him and make him a 'case'. (p. 192)

This means that the certification we gain as a result of our subjection to processes of examination confers an identity—we become socially visible as possessing the qualities the examination seeks to establish. We thus become socially visible in the social roles—the subject positions—available within the discourse. Examples of this process will be given below, in relation to the certification of immigrant health professionals, and in the identities conferred within policy-related assessment frameworks, which dominate contemporary language education. Accepting such a theory of social context would bring about significant changes in the way we think about our tests and our role as language testers. It would impel us to envisage the testee not only as an individual with cognitive attributes defined in terms of the theorized features of the immediate situation of use, but as a social being whose subjectivity is a function of subject *positions* realized in the administration and use, the very existence, of the test itself. From this perspective, tests become technologies of subjectivity. They act as mechanisms both for the definition of subject positions and for the recognition of subjects. Tests create the identities they measure.

Subject positions are articulated within discourses. For example, the subjectivities defined within the discourse of modern clinical practice are the topic of Foucault's *The Archaeology of Knowledge* (1972 [1969]). Foucault defines clinical medicine as a discourse:

Clinical medicine must ... be regarded as the establishment of a relation, in medical discourse, between a number of distinct elements, some of which concern[ed] the status of doctors, others the institutional and technical site from which they speak, others their position as subjects perceiving, observing, describing, teaching, etc. It can be said that this relation between different elements ... is effected by clinical discourse: it is this, as a practice, that establishes between them all a system of relations ... (p. 50)

Foucault discusses each of the elements in the system of relations that constitutes a discourse: speaker status, site of speaking, and subject positions of speakers. All of them are constituted in part or in whole in language. This means that the question of who is accorded the right to speak in such a discourse, and from where, and occupying which subject position, is obviously of crucial relevance to understanding the function of the language tests that guard access to this right.

On the status of the speaker, Foucault writes:

First question: who is speaking? Who ... is accorded the right to use this sort of language? Who is qualified to do so? Who derives from it his own special quality, his prestige ... ? What is the status of the individuals

who — alone — have the right . . . to proffer such a discourse? . . . Medical statements cannot come from anybody . . . (p. 51)

The status of speaker in medical discourses is rigorously guarded: as a result, immigrant health professionals, who have gained the right to speak in their own culture, need to go through a process of recognition before that status is again conferred on them in the new cultural setting. Language examinations are a typical aspect of such processes.

Foucault goes on to consider the subject positions defined within medical discourse:

The positions of the subject are also defined by the situation that it is possible for him to occupy . . . : according to a certain grid of explicit or implicit interrogations, he is the questioning subject and, according to a certain programme of information, he is the listening subject; according to a table of characteristic features, he is the seeing subject, and, according to a descriptive type, the observing subject: . . . To these perceptual situations should be added the positions that the subject can occupy in the information networks (in theoretical teaching or in hospital training; in the system of oral communication or of written document: as emitter and receiver of observations, case-histories, statistical data, general theoretical propositions, projects, and decisions). (p. 52)

These subject positions are in contemporary language testing practice the basis for the content of relevant language tests. What makes his analysis different from the job analysis stage in the development of specific purpose language tests, which it superficially resembles, is that such subjectivities are implicated in relations of power; thus Foucault's analysis is not an exercise in pragmatism but a social critique.

Foucault's notion of discourse clarifies for us the socially contextual function of a test of English for the certification of immigrant health professionals, such as the OET (McNamara, 1996). This Australian test has existed in one form or another for over 30 years, as part of the process required by the professional registration authorities for the recognition of the professional credentials of immigrant health professionals. It was reformed in the mid 1980s to reflect the communicative demands of the workplace, and now it contains assessments in the four skills of listening, speaking, reading, and writing based on tasks that are characteristic of the workplaces of the various health professionals for whom the test is designed. From a Foucauldian perspective, the function of the OET is clear: it is a technology for recognizing (that is, publicly confirming and certifying) individuals as suitable objects of such subject positions. In this sense, the OET, like other tests, is a test of identity. Like all tests, it represents a site of social recognition and control.

What advantages are there for us as language testing professionals seriously engaging with a theory of the social context of testing?

First, we can avoid being merely naïve players in this discursively constructed world. With appropriate intellectual and analytical tools we are enabled to become aware of the roles that tests will play in the operation of power and of systems of social control. We will be less inclined to shelter in the impersonality of the purely technical aspects of our work. We need critical self-awareness in order for us first to become aware of and then to decide whether to accept or to resist our own subject positions in the system of social control in which tests play such a part. LTRC as a conference will enable us to reflect on our subject positions — our identities — as language testers, rather than simply reinforcing them.

Second, we will be able to understand how the emphasis on the cognitive and psychological in measurement is part of the means through which the social construction of subjectivity, and hence the social function of the test, is rendered invisible. The point about subjectivity is well made by Butler (1990) in her discussion of the performative nature of gendered identity:

If the 'cause' of desire, gesture and act can be localized within the 'self' of the actor, then the political regulatory and disciplinary practices which produce that ostensibly coherent gender are effectively displaced from view. The displacement of a political and discursive origin of gender identity onto a psychological 'core' precludes an analysis of the political constitution of the gendered subject. (p. 136)

Third, attention to social theory will help break down the isolation of language testers from areas of the humanities in which social theoretical awareness is articulated, which I see as currently one of the most vulnerable aspects of our field. Without critical awareness ourselves, we will be ill-equipped to engage with the challenges presented by our critics.

Fourth, a theorization of the social context of testing will help us understand the discourses in which the demands of our sponsors and clients are shaped. This will better prepare us to expect, to recognize, and to deal as constructively as we can with various kinds of pressures on tests as a result of their function as sites of social control. For example, the OET is subject to relatively abrupt changes in policy, both in who will be required to take it and at what level of performance permission to practice will be granted. This is because the test acts as a site for the insertion of the power of competing discourses, of government responsibility for provision of medical services in settings of scarcity of medical resources, such as in rural areas, of patients' rights, of multiculturalism, and so on. Other examples of test as sites of social control are the dictating of test constructs as a function of government policy, as in the mandating of levels on the Common European Framework of Reference

for Languages (Council of Europe, 2001; Fulcher, 2004b) or the imposition of testing regimes such as the No Child Left Behind Policy (Kunnan, 2005b; Byrnes, 2005). We should also expect that those in power will want to retain control of the setting of cut scores in the interests of policy objectives. A clear example recently has been the role of language tests in current discourses on immigration and citizenship. The legislation on immigration and citizenship in several European countries makes reference to levels on the Common European Framework of Reference for Languages (Council of Europe, 2001). There are six levels, A1 and 2, B1 and 2, and C1 and 2, spanning the range of language proficiency from beginner to fluent, native-like command; A1 is the lowest level. The levels required under this legislation differ by country, depending on the politics of the setting. Austria has recently increased from A1 to A2 the language requirement that immigrants need to demonstrate within 18 months of residency, or face loss of residence rights (König and Perchinig, 2005); the Netherlands requires new migrants to reach A2 (Marinelli, 2005); Finland requires A2 for members of Finnish ethnic communities in Russia wishing to reside in Finland; recent changes to both the German and Danish citizenship laws require B1. The politics and cultural values of the setting in each case appear to be the decisive factor, as it is hard to see what justification there might be for the different levels in terms of the functional requirements of citizenship, which are presumably equivalent across countries.

Finally, engagement with the theory of the social context represents a response to one of the requirements of test validation as expressed by Messick (1989), that the values implicit in test constructs be investigated and articulated. These values can be understood in social and policy-related terms and can be revealed by considering the discourses within which language tests are located and have their meaning. The extent to which individual test developers should be required to demonstrate such a consciousness is obviously highly debatable, and here we come up against the limits of current validity theory as represented by Messick and Bachman, in that, lacking an adequate social theory, it is not adequate to deal with the perspective made available within a social theoretical orientation (McNamara and Roever, 2006). But even if test developers do not carry out the kind of research being advocated here, there is clearly a need for this research to be carried out within language testing as a research area more broadly conceived.

For all the reasons cited, theories of the social context of language testing should be the site of urgent exploration. We obviously have a long way to go if we are to develop such a tradition of addressing the social context of language tests at our conferences and in our research. But the challenge is both intellectually exciting and necessary for informing our awareness and our practice.

This page intentionally left blank