

Language testing within policy contexts: Conceptual and instrumental challenges

Tim McNamara
The University of Melbourne

The establishment in 1990 of what came to be known as the Language Testing Research Centre (LTRC) at the University of Melbourne was made possible by Australian government initiatives in the field of language policy, specifically the adoption of the National Policy on Languages (Lo Bianco 1987). Seed funding was made available for the establishment of a number of research centres in applied linguistics to support the implementation of the policy. Despite the policy location of the work of the Centre, an awareness among the language testing researchers working there of its intrinsically policy-related character has taken a long time to emerge. This has been perhaps due to the training in applied linguistics, which has emphasised cognitive and pedagogic approaches to language learning. I remember being shocked when the founder of the Melbourne Applied Linguistics program, Dr Terry Quinn, mentor to both myself and Joe Lo Bianco, said to me once about a well-known researcher in language testing: 'Imagine, he doesn't understand that he is engaged in essentially a political activity'. I was already a few years into my own language testing career, and winced in self-recognition. Now, thirty years after the founding of the LTRC, the field has become more and more conscious of its policy and political character, and the four papers in this issue represent significant advances in our thinking in this area.

The papers demonstrate the character of applied linguistics more generally, as both instrumental and conceptual (Elder, this issue). Applied linguistics has always struggled with getting the balance right between these two imperatives. The founding Director of the LTRC, Alan Davies, was oriented to social issues as well as learning issues in his applied linguistics work, but insisted on the centrality of the conceptual work of the Centre, which has given it an enduring strength. But the Centre has also needed to be practically oriented as it is an entirely self-funded research Centre, and has relied for its survival on delivering the instrumental goods to its clients. How to build conceptual research into that instrumental funding environment? The four papers show how innovative conceptual work can be grounded in the practical research projects which the Centre undertakes.

The paper by Frost points out the conceptual contradictions between current validity theory and the reality of the bureaucratic use of test scores, specifically within the Australian immigration process, and more generally. Her critique is deeply radical, as

it is not clear how validity theory as currently understood can cope with this contradiction; it exposes the liberal assumptions of validity theory, inadequate to account for the actual uses of tests in policy settings. Her paper thus questions the adequacy of even the more policy-oriented approaches to validity in the work of Chalhoub-Deville (2009; Chalhoub-Deville and O'Sullivan, 2021), cited too in Elder's paper. The paper by Macqueen et al. further develops our understanding of the way bureaucracies interpret test score meaning. The paper adopts the notion of 'trust' from the political science literature to show the power and attraction of tests and test scores in the work of professional and academic bureaucracies.

In the case of each of these papers, the authors have gone well beyond the established literature in our field to explore the usefulness of concepts from sociology, political science, discourse studies, and so on. Both Elder and Knoch in their papers also embrace theoretical concepts from other fields, notably policy studies and language policy in particular. Elder gives a useful summary of relevant work on policy discourse by Joe Lo Bianco. It is ironic that we have not explored the usefulness of his work to ours earlier, given that he was responsible for the developments that led to the establishment of the Centre, and the fact that we have worked together in the same University for a number of years. The need to engage with discourses beyond those that are familiar within language testing is clear from all four papers. This also has implications for the training of language testing researchers, a topic which Knoch in her paper addresses.

The arguments of all four papers are in turn grounded in empirical work carried out by Centre staff, some of it dating from the earliest days of the Centre. This is particularly true in the case of the papers by Elder and Knoch. Knoch recounts three illustrative episodes in which she and other Centre staff were called on by external authorities to provide expert advice on language testing matters. She shows both how complex a process the provision of this advice was, particularly in terms of fitting it to its audience, the patchy success with which the expert advice was taken up, and the likely reasons for this. Elder produces a deeply impressive list of the projects undertaken by the Centre in relation to the teaching of languages other than English (LOTE) in Australian schools and universities. She revisits key examples from the projects and reinterprets them using a language policy lens, which helps us understand the degree to which their findings were taken up in particular policy settings, or quietly shelved. The case studies offer a fascinating picture of the complex forces operating within contexts in which language testing expertise is made available or is solicited.

In general, the papers in this issue provide a snapshot of how our policy-centred field is slowly awakening to a self-consciousness of its character, and articulating the dilemmas and challenges that this new awareness brings. It is remarkable that it has

taken our field so long to reach this self-awareness. Conceptually it is a necessary and important advance, long overdue. More specifically, and more fundamentally, the experiences documented in the articles pose huge conceptual challenges within the fundamental theoretical basis of our field, validity theory. Practically, and instrumentally, reflection on the collective experience of those engaging explicitly with policy contexts may suggest some useful ways forward, even if the theoretical issues remain for the moment intractable.

References

- Chalhoub-Deville, M. (2009). The intersection of test impact, validation, and educational reform policy. *Annual Review of Applied Linguistics, 29*, 118-131. <https://doi.org/10.1017/S0267190509090102>
- Chalhoub-Deville, M. & O'Sullivan, B. (2021). *Validity: theoretical development and integrated arguments*. Equinox.
- Lo Bianco, J. (1987). *National policy on languages*. Australian Government Publishing Service.