

The challenges of providing expert advice in policy contexts

Ute Knoch
University of Melbourne

As academics we surrender our expertise to a conversation in which we can shape and influence an outcome, but cannot determine it (Lo Bianco, 2019, p.99)

Language testers can have various roles in relation to the use of language tests and policy. One such role may be to provide expert advice in policy formation or policy review contexts. Such instances are often not documented systematically, as advice may be provided in informal or closed meetings, or confidential documents, which are not available to the public. The challenges associated with providing such advice are also rarely recorded. In this paper, I describe three instances in which the Language Testing Research Centre (LTRC) at the University of Melbourne was invited to provide external policy advice. I specifically reflect on what prompted the invitations, what advice we were asked to provide, the complexity of providing advice in each instance, and whether the advice was taken up by the policy makers. The paper reflects on some of the common threads and concludes with implications for training new language testers.

Keywords: policy advice, policy formation, invited expertise

Introduction

Language testers can take a variety of roles in policy formation and evaluation – they can be insiders and be directly involved in shaping the content of a policy or a project that implements or evaluates a policy, or they can be an outsider looking in and providing an unsolicited critique of the policy. Finally, and the focus of this paper, is a third situation where language testers are invited to provide academic expertise to policy makers in various settings. This topic is largely neglected in the language testing literature, and, as I will argue later in the paper, in the training of language testers. The paper argues that drawing on literature in public policy can provide a useful backdrop to understanding policy formation situations in applied linguistics, and why or why not advice is taken up in such contexts.

Literature review

According to Lo Bianco (2019), policies exist in an 'environment of agitation for additional, replacing, or supplementing action' and for this reason policy is recognized as 'part discourse and part text' (p. 91). The text exists largely under the control of officials responsible for the policy and under the management of politicians; the discourse on the other hand is available to the wider community which is able to influence policy texts through direct action, through media, research and a range of other means (Lo Bianco, 2019, p.92). Academic researchers, according to Lo Bianco, are able to contribute in the policy formation or revision process with the ordinary voices of citizens and again, as an 'expert voice' (p. 92) which will carry extra power. Bacchi (2009) points out that policy makers usually find the citizen voice more persuasive than the expert voice because the citizen voice is directly linked to votes, indicating that the power of expert advice may be limited in certain settings (see also Fischer & Gottweis, 2013).

When language testers do get invited to provide expert advice to inform policy decisions, what exactly happens in such contexts is very rarely documented publicly; nor are the reports they write usually publicly available. Researchers often also have to sign confidentiality agreements when they are engaged for their work. This makes it difficult for young applied linguists and language testers to learn about this type of work from their more experienced peers.

Much of the literature on policy and policy-making in language assessment and applied linguistics is inward-looking. The literature on public policy provides a theoretical backdrop that is worthy of further exploration. Jann and Wegrich (2007) proposed the idea of the policy cycle which comprises agenda-setting, policy formation, decision-making and implementation. The policy cycle has been criticised as being a simplification of the complex reality of policy making and the multi-faceted nature of policies involving complex links to previous as well as other, potentially competing policies. Policy processes also rarely feature clear beginnings and endings; instead, the stages are 'entangled in an ongoing process' (Jann & Wegrich, 2007, p.44). Expert advice may be requested at several of these stages and its scope and effectiveness may be compromised by the particular stage of a policy, as well as the 'wickedness' of the problem (Fischer & Gottweis, 2013, p.429) and the complex competing views of governmental officials, politicians and members of the public (in the case of larger political contexts), or the web of stakeholders in other contexts. Kim et al. (2018), concerned with how research findings are taken up in global health policy contexts, developed a research utilisation framework, comprised of four phases. The circular model includes a number of key decision points in which research knowledge can be translated and adopted into policy. It also includes

groups of key stakeholders (called key actors) taking on different roles within the policy cycle. While expert advice provided by applied linguists is not always only based on research findings, this model is likely to be applicable to such contexts. The authors also argue that the use of evidence may be of a political nature, as decisions are often made based on relationships and political dynamics, which may be difficult to understand for researchers.

Expertise in applied linguistics and language testing

Grabe (2010) describes how applied linguistics has changed over its history claiming that it is 'the language-based problems in the world that drive applied linguistics' (p. 8). He lists various problems that applied linguistics might address (including language policy problems) and notes that the attempt to solve these problems often leads applied linguists to draw on knowledge from a range of fields apart from linguistics. Chapelle (2013), however, notes a mismatch between the aspirations of the coverage of applied linguistics (defined in her work as any domain of activity where an understanding of language might inform practice or policy), and what many applied linguists actually do. She proposes that one explanation for this may be that many applied linguists come from a narrow background and are often focussed on just one sub-area (such as language teaching or linguistics) and are therefore not able to adequately offer expertise to solve real world problems.

Wray and Wallace (2015) explored what it means to have expertise in applied linguistics through interviewing well-respected and experienced applied linguists. Key results from their study show that these scholars work at the boundaries of their knowledge, are known to have depth and breadth of knowledge and the capacity to communicate ideas effectively. They also acknowledge the role of a good quality research environment to develop knowledge, skills and confidence. Experts in their study also discussed the importance of knowing and acknowledging their own limitations in knowledge and being able to collaborate with others.

Focussing more specifically on language testing as a sub-discipline of applied linguistics, it is worthwhile reflecting on the training of language testers. Most language testing specialists study in applied linguistics programs at Masters level, where they may be exposed to the subject language testing/assessment for the first time. This subject usually provides a good overview of assessment of the language skills (writing, speaking, listening, reading), of key theories underpinning language assessments, psychometrics, concepts of validity and validation, and increasingly in recent years also includes a focus on the ethics of language testing and test consequences (Brown & Bailey, 2008; Jin, 2010). Those language testers who go on to pursue a PhD in language testing, are typically,

depending on the university, required to conduct a larger research project for their dissertation and may also need to engage in further course work. No training course for language testers, to my knowledge, systematically addresses how to best engage with policy makers. This is something that is often only learned by experience by those who are engaged in such situations.

Interest in language assessment literacy (Inbar-Lourie, 2008; Kremmel & Harding, 2020; Taylor, 2009) of various groups of professionals (e.g., teachers, test providers/developers, language testing researchers) and non-specialists (Pill & Harding, 2013) has grown in recent years and has resulted in the modelling of aspects of this knowledge that might be needed by these stakeholder groups. The components included in these models have slightly changed over time, with Kremmel and Harding (2020) being the first to call one factor that emerged from their questionnaire ‘assessment policy and local practices’. Interestingly, none of the questionnaire items that loaded onto this factor were related to directly dealing with policy-makers – questions were more concerned with policy makers’ need for knowledge of whether assessments are suitable to specific contexts, or what assessments can be used for, rather than with the ability of language testers to work effectively with policy-makers.

Another possible source of training for language testers is written publications. However, to my knowledge, there is nothing available that may help train testers in providing expert advice. One document that provides at least some guiding principles, is the International Language Testing Association (ILTA) Code of Ethics (International Language Testing Association, 2000; 2018). Principle 7 of the Code of Ethics, which is the most relevant to this paper, and selected annotations are reproduced below:

Principle 7

Language testers in their societal roles shall strive to improve the quality of language testing, assessment and teaching services, promote the just allocation of those services and contribute to the education of society regarding language learning and language proficiency.

Annotation

- Language testers shall be prepared to work with advisory, statutory, voluntary and commercial bodies that have a role in the provision of

language testing services.

- Language testers shall be prepared to interpret and disseminate relevant scientific information and established professional opinions to society. In so doing, language testers shall clarify their status as either spokespersons for a recognised professional body or not. If the views expressed are contrary to those generally held, they shall so indicate.
- It is reasonable for language testers to make scientifically substantiated contributions to public debate on sensitive socio-political issues, such as race, disadvantage and child rearing.
- Language testers shall differentiate between their role as educators based on professional knowledge and their role as citizens.
- In fulfilling their responsibilities under this principle, language testers shall take care to avoid self-promotion and the denigration of colleagues.
- Language testers shall make clear that they do not claim (and are not seen to claim) that they alone possess all the relevant knowledge.

As can be seen, the Code of Ethics sets out some useful aspirations for language testers engaged in public or policy debate and advice. However, due to the lack of training and guidance provided to language testers (and probably applied linguists more broadly) on working with policy-makers, and the lack of written resources to draw on for those trying to familiarise themselves with how best to approach such work, there is very little those new to the profession can do to prepare themselves for such work. These lacunae provided the impetus for this paper, which aims to document and reflect on cases of policy-advice work undertaken by staff in the Language Testing Research Centre (LTRC) at the University of Melbourne.

The current paper

For the purpose of this paper, I selected three very different scenarios in which staff at the LTRC were invited to provide policy advice in policy-formation or policy-evaluation contexts. In particular, I describe four aspects of each situation:

- (1) How we came to be invited to provide expertise,
- (2) What advice/expertise we were asked to provide,
- (3) What complexities we encountered when providing the advice,
- (4) Whether the advice was taken up.

It is important to note, at this point, that my reflections on each of these cases are based on personal experience alone and may not represent the experiences of colleagues.

The three policy-advice cases described in this paper were selected to represent a range of policy contexts, the first being a federal government policy, the second a policy prepared by a professional registration authority, and the third a local university policy. Apart from the level of the policy contexts, the three cases also differed along some of the other dimensions we list above. The three contexts were selected because the policy advice situation was concluded or nearly concluded by the time of writing this paper.

Case 1: Introduction of the Australian Citizenship Bill 2017

In 2017, the Australian government (which at the time consisted of a coalition of two conservative parties), proposed the introduction of amendments to the Australian citizenship legislation (The Australian Citizenship Legislation Amendment - Strengthening the Requirements for Australian Citizenship and Other Measures - Bill 2017). Relevant to our action, the amendments focussed on two aspects: (1) the introduction of a new English language requirement in the form of a standardised large-scale English language test (the IELTS test, with Band 6 set as the passing level), and (2) a 'strengthening' of the already existing knowledge of society (KoS) test to include a focus on Australian values. The bill was initially introduced into the Lower House. At the time, submissions were invited by the Department of Immigration and Border Protection - the LTRC made a submission in May 2017. The submissions were not made public, departing from usual practice. The bill was then debated in the Lower House and passed. It was then referred to the Upper House, where the government did not hold the majority. At this point, the proposed bill received substantial media attention, in particular the fact that the proposed English language proficiency test was set at Band 6, which required 'university-level English'. At the time, several LTRC staff members were interviewed on various media platforms about details of the bill.

The proposed bill then became the subject of a senate enquiry, and submissions were again invited. The LTRC made a submission in which we focussed on the following themes.

- 1) We questioned the need for a separate English test, citing research evidence showing that immigrants could contribute effectively to Australian society with relatively low levels of English.
- 2) We claimed that the academic nature and the literacy demands of IELTS (including the General Training module) rendered it unsuitable for use in this context.

- 3) We pointed to the four skills tested by IELTS and proposed that, if a stand-alone test were deemed necessary, a functional test of conversational English would be more appropriate.
- 4) We claimed that the requirement of IELTS 6 was unwarranted, being higher than that required in other countries and beyond the reach of many current citizens with limited English literacy.
- 5) We argued that if a stand-alone English test were introduced, immigrants who had already demonstrated the language proficiency required for permanent residency should be exempt.
- 6) We opposed the proposal to extend the Australian KoS test (which was already linguistically demanding in its current form).
- 7) We suggested that the government should focus on facilitating easier access to English tuition to migrants.

Many other groups and individuals also made submissions (635 written submissions were received) representing a range of organisations, including legal experts, teacher associations, human rights groups, multicultural organisations, and refugee support groups, with nearly all submissions opposing the bill (Elder et al., 2019; Macqueen & Ryan, 2019). Following our submission to the senate enquiry, several hearings for the inquiry were set up, and a representative of staff from the LTRC was invited to attend. Associate Professor Cathie Elder appeared at the Melbourne hearing. The proposed bill was later voted down in the Senate.

In the following, I reflect on the four areas of focus mentioned above.

(1) How we came to be invited to provide expertise

As described above, we were invited to provide expertise based on our submission to the Senate enquiry, which commented on a range of issues around language testing and language proficiency in relation to the bill. Our submission positioned us as experts in the field. Unrelated to the Senate enquiry, we were also contacted by several media outlets to comment on the proposed changes, including the construct of the IELTS test, how the general and academic modules of the test differed, and the level that was required.

(2) What advice/expertise we were asked to provide

In this section, we will comment on the advice that we were asked to provide at the senate enquiry, as this was the only context in which policy-makers directly asked questions. We have already commented on the content of our written submission above. During the

Senate enquiry, Cathie Elder was asked to speak to the senate submission, and commented on the language demands of the existing KoS test, the differences between the IELTS academic and general modules, the meaning of a IELTS level 6, and the relevance of the test to the every-day life domain of aspiring citizens. She was also asked to provide follow-up information on notice in response to her suggestion that there might be other, more productive means of building English language proficiency than the introduction of a standardised testing requirement. She subsequently provided follow-up documentation about possible English language study pathways to citizenship as alternatives to the proposed test pathway.

While not in the business of providing policy advice, the Australian media picked up on community concerns about the minimum level of English to be required of aspiring citizens and were interested in the meaning of the score level, and what this represented. They started referring to IELTS 6 as 'university-level English'. Whether IELTS 6 represents university-level English became a point of contention, as the minister in charge of the proposed bill denied that this was the case in an interview, and media channels then set out to 'fact-check' this particular point.

(3) What complexities we encountered when providing the advice

One lesson we learned from our experience of public intervention was that we may need to be prepared to mount an argument with supporting evidence within a very limited timeframe. It was challenging to assemble relevant evidence for points we were arguing, for example, about the impact of introducing a large-scale standardised test on vulnerable populations. We only subsequently located data from other countries which would have assisted us in making this case.

A further lesson we learned from our experience was the need to have a better understanding of the priorities of politicians as well as the pressures they are under from voters in their electorates. In hindsight, it may have been more useful, rather than simply raising objections to the proposed testing requirement, we should also have proposed alternatives (as Cathie Elder attempted to do later at the request of the chair of the Senate enquiry). Proposing an alternative course of action that policy makers can adopt may mitigate the loss of face they experience when their initial proposals are defeated. Equally, we learned through the experience of the Senate enquiry as well as our media opportunities that it is really important to be skilled at presenting very simple messages to outsiders in ways that catch their attention and appeal to their interests.

(4) Whether the advice was taken up

The proposed bill was eventually voted down in the Senate. It is difficult to attribute this outcome to any specific action, in particular given that the conservative government at the time did not have the majority in the Senate and therefore the bill may have been voted down regardless of any input from the inquiry. As mentioned above, there was a lively debate amongst parliamentarians around the bill before it passed the lower house, and the subsequent Senate Inquiry received 635 written submissions from diverse individuals and groups – our submission was just one among many inputs. There is, however, some evidence that our particular contribution was noticed during the face-to-face hearing of the Senate Inquiry and the final report prepared by the Senate’s Legal and Constitutional Affairs Legislation Committee. In the Melbourne hearing, the Committee chair requested further information on notice in response to our suggestion that there may be more productive ways of building English proficiency than the introduction of a standardised English testing requirement. Furthermore, in the Committee’s final report, there is evidence that points to uptake of our evidence, citing information presented by Cathie Elder at the Senate Hearing, as well as elements from the written submission by the LTRC (Knoch et al., 2017).

Case 2: Professional Registration Authority

The second case involves a project the LTRC completed for an Australian professional registration authority. One of this authority’s tasks is to ensure that overseas-qualified professionals meet the various registration requirements in Australia, including the English language skills requirements (see Macqueen et al., and Frost, this issue). As part of this work, they engage in a regular review of the English language skills required for registration in Australia. The last such review took place some years ago. To commence the next cycle of review, the authority issued a request for work through a competitive tender process, which the LTRC responded to and was subsequently awarded. The scope of work focussed on the two main strands of demonstrating English language skills, namely, the test pathway, and the non-test pathway.

The previous standards set out a range of possible pathways for registration. Although there are some slight variations in the requirements, for the majority of professions, overseas applicants can either present an English language test score from one of four approved tests (IELTS, PTE, TOEFL, OET) at the required minimum level, or they can register through non-test pathways where applicants prove English language proficiency through showing education in English-speaking schools or higher education institutions.

In what follows, I again consider the four areas of focus for the description of each case.

(1) How we became invited to provide expertise

We were approached by representatives from the authority at the recommendation of an academic in our university and were invited to respond to the request for proposals. We were subsequently chosen to undertake the work.

(2) What advice/expertise we were asked to provide

The request for proposal asked us to provide information on the language tests that registration authorities in other countries accept for registration and whether any research is available that supports the minimum levels used for entry into the profession. Regarding the non-test pathway, we were asked to investigate what non-test pathways are used by registration authorities in other countries.

Apart from the issues requested by the agency, we also provided a description of key differences between general academic and occupation-specific tests, and how well these two types of tests match the communicative domain, key issues in validation in language testing, various issues relating to standard-setting for professional purposes and how standards on different tests can be compared.

To inform our discussion of the non-test pathway, we cited literature on language development of students studying in English-medium universities. This issue was relevant as there is an assumption that students' language proficiency develops at university, and therefore students increase their language proficiency from entry into the university to exit.

(3) What complexities we encountered when providing the advice

There were two particular areas that we found challenging to respond to, both related to the non-test pathway. Firstly, it was difficult to comment on non-English speaking background students' language development in English-medium universities, as this issue is complex to investigate, and the studies that are available (Gan et al., 2015; Humphreys et al., 2012; Knoch et al., 2015; Knoch et al., 2014; O'Loughlin & Arkoudis, 2009; Storch & Hill, 2008) do not provide sufficient information to provide sound advice to the authority to inform non-test pathways.

Secondly, although not directly mentioned in the request for proposal, the authority was interested in language use/language development of students who enrol in online degrees. Unfortunately, we could not find any information on this in the research literature, and therefore found it hard to comment.

(4) Whether the advice was taken up

This question is difficult to answer at this point. The authority is currently in the process of making some changes to the registration requirements, and these proposed changes will then go out for consultation. The changes we have been told about are small, and involve the additional acceptance of two tests, as well as some minor changes to the wording of the non-test pathway. Many of the recommendations we made appear to have not been taken up in these changes, and the reasons for this are highly likely to be practical. Policy makers are answerable to multiple constituencies and our advice is only one source of information amongst many sources of input for their deliberations.

Case 3: University entry requirement (change during COVID)

The administration of many high-stakes language tests was disrupted in parts of 2020 (and beyond) due to the COVID-19 virus. This forced institutions who relied on English language test results for admissions to adapt and think of alternative solutions to gather evidence from prospective students. The third case chosen for this paper relates to a policy decision made by a large Australian university to deal with the problem that students were not able to produce evidence to satisfy the university's English language requirements due to the suspension of most large-scale English language tests in 2020. This was particularly an issue in China, the country of origin of the majority of overseas students at this university, where test centres of almost all the major tests were closed or were operating at reduced capacity to ensure physical distancing. Due to security and validity concerns, the university was reluctant to accept the various at-home online tests that were offered as alternatives as the only evidence of English language proficiency for entry into the university. To provide another option for students wishing to study at the university, the university's direct entry English language school put forward a proposal to provide a two-week course for students who had taken a suite of English language at-home tests. The proposal was that this would culminate in an assessment offered online by the English language school which would be seen as further evidence of English language proficiency over and beyond the at-home English language tests. The rationale for offering a two-week course (rather than asking students to take the assessments designed by the English language centre directly), was that students should be given the opportunity to familiarise themselves with the assessments offered, and to be provided with the opportunity to practice responses to these tasks to ensure fair opportunities on the final assessments.

Below, I consider the four areas of focus:

(1) How we came to be invited to provide expertise

The university's admissions committee approached the LTRC and asked for recommendations on possible cut-scores for the at-home screening tests which would provide the first evidence and would form the basis for admission into the two-week course offered by the English language school.

(2) What advice/expertise we were asked to provide

We were tasked to take on a very limited role – comment on potential cut-scores for screening tests – rather than provide a fuller evaluation of the entire policy, the assessments used for both screening purposes and as final course assessments. Rather than only commenting on the cut-scores, the LTRC decided to provide a review of the full proposal, including the suitability of the course offered, the final assessments, and the possible screening tests to be accepted (as well as the cut-scores on these), together with a short review of each of these tests. The report sent to the university committee also included a number of recommendations, including ongoing monitoring of students who have entered through this pathway to ensure adequate support is provided to students once at university. In terms of the at-home English language tests which would provide the first screening mechanism for students before entering the course, we recommended that cut-scores be set at an equivalent of the minimum English language entry scores at the time, and that only at-home tests be accepted that measure academic English language skills rather than general English language skills.

(3) What complexities we encountered when providing the advice

No major complexities were encountered, although we did have some problems finding key information about some of the at-home online English language tests that were proposed as possible options. Not all test providers had posted all the necessary information on their test websites, probably because many of these tests were made available at short notice, and key validity research underpinning the tests was not yet available. Some of the details of test security (e.g., proctoring) and availability (e.g., whether a test was available in China) were also not always clearly described on the test websites.

(4) Whether the advice was taken up

The university leadership appreciated the report, but immediately rejected the advice we had given on setting the cut-scores on the Duolingo test (which was one of the possible screening tests under consideration by the university, and therefore included in our proposal). The equivalent to the current English language entry scores was rejected as being too low, as other universities had set these much higher. The university leadership,

however, disregarded in this matter that other universities' cut-scores on at-home tests like the Duolingo were set higher because they served as the *only* evidence of English language proficiency, while the proposal here was to merely use this as the first screening instrument to enter the two-week course.

The entire proposal put forward by the English language centre was later approved (it is not clear whether this included the use of the Duolingo English Test, and if so, what minimum scores were selected), but the course never went ahead as no students enrolled.

Discussion

This paper described three policy-advice situations in which language testers were invited to provide expert advice to inform policy-making or policy-amendment processes. I have detailed above how these cases differed along four dimensions: (1) how we became involved in providing advice, (2) what advice we were asked to provide or what advice we provided, (3) any complexities that we experienced in providing this advice, and (4) whether the advice was taken up.

The road to providing advice in the three cases differed. In the case of the citizenship bill, we responded to an open call for submissions to a public inquiry and were then selected as 'experts' based on that submission to attend a public hearing of the Senate Inquiry. In the second case, we prepared a submission to an open call for tenders and were selected to do the work. However, we were approached by representatives of the authority and asked to respond to the tender – it is not clear what the process was that led to this in the first instance; nor do we know whether other groups responded to the tender. In the final case, we were directly contacted by a representative of the Office of the Provost with very specific questions about cut-scores. The recognised expertise of the LTRC, both internal and external to the university, is likely to have led us being approached for advice in the second and third case, whereas in the first case we took the initiative in positioning ourselves as experts to politicians and policy makers. It can be seen that all these avenues differed, and were often a matter of serendipity. This shows that even if specialist expertise is available, there may be no opportunity to provide it.

The expertise that we provided or were asked to provide differed in all of the cases, although some common themes emerged. In all cases we commented on the nature of the testing constructs and their relevance to the domain in question. In all cases, we provided advice on cut-scores, either on those already chosen by others (Case 1), possible ways of setting cut-scores or what cut-scores other, similar organisations were using (Case 2), or on cut-scores we thought were reasonable in a first stage of a two-stage testing program

(Case 3). In Cases 1 and 2, we also provided information that went beyond language testing expertise about expected language learning progressions of various (vulnerable) populations to discuss possible test impact (Case 1) or non-test pathways (Case 2). These latter cases show that training as a language tester is necessary but by no means sufficient in such policy-advice cases, and concurs with the finding by Wray and Wallace (2015) and the claim by Chapelle (2013) that depth and breadth of knowledge is needed in such situations.

The stage of the policy cycle may impact the type of expertise requested and the likely uptake of this expertise. Case 2 was at a later stage of the policy cycle (Jann & Wegrich, 2007) than the other two cases, with the aim of the work to inform the possible revision/evaluation of an already established policy. Cases 1 and 3, on the other hand, were in the early stages of the development of new policies, possibly influencing uptake of advice but also the number of stakeholders potentially involved in providing input to the policies.

The three cases collectively show that the provision of expertise may be limited due to a number of reasons. Firstly, we found in the first two cases that an insufficient research base was available to provide very detailed advice on language learning progressions of certain populations. Where research was available, we did not have that information at hand in real time (Case 1). As Berthele (2019) argues, applied linguists who act as experts in language policy-making situations need to be aware of their own ignorance and personal biases to ensure that advice that is provided is not based on an insufficient research base, 'bad science, vague theories and studies biased towards confirming our beliefs' (p.8). It is of course also important to make the point that policy questions/problems are generally much broader than research questions in academic work, so even a relevant line of research is not necessarily adequate as a basis for advice on real, practical policy matters. Policy decisions require a substantial research agenda, which is often too expensive or too late for the decision timeframe.

Finally, the uptake of our expertise varied across the projects. In Case 1, some aspects of our expertise were reflected in the final report prepared by the Senate's Legal and Constitutional Affairs Legislation Committee. Uptake of advice in political contexts is of course mediated through a range of factors. Fischer and Gottweis (2013) argue that an empirical approach to policy enquiry is insufficient. They argue that language and argumentation are key to policy analysis (and hence also policy-advice) and is more suitable to today's 'messy policy problems' (p. 428), in particular in complex political contexts 'to capture the typically heterogenous, interconnected, often contradictory, and increasingly globalized character of these issues' (p. 429). Uptake of policy advice

requires ‘understanding of human behaviour as culturally influenced, communicatively developed, emotionally grounded and socially or politically motivated’ (p. 429) (see also Kim et al., 2018). In Case 2, it is yet unclear whether any of our advice was taken up, but the proposed changes that have recently been put out for consultation indicate that the majority of our recommendations were not taken forward, for whatever reason this may be. And in the third case, university admissions testing, our advice was also rejected in favour of practice from another university. In Cases 2 and 3, our academic recommendations were rejected because the policy-makers were more influenced by the policies of other universities in relation to cut-scores than by our recommendations, a situation that Furedi (2011) and Deygers and Malone (2019) likened to policy-shopping. From these examples, it can be seen that there are clear limits to what expert policy advice can achieve, a fact also noted by Lo Bianco (2019).

Taken together, all these three cases provide an insight into the complexities of providing expert advice and the limited influence we, as language testers, may have on policy outcomes. Elder (this issue) calls for more ‘policy-responsible’ language testers and testing practices. The three case studies presented in this paper illustrate the limited influence language testers may have on the outcome of policy-formation and/or revision if their role is to provide advice, rather than being involved in a fully funded project, such as the cases described in Elder’s paper. Even the opportunity of being invited into the policy context to provide advice is neither easy nor straight-forward, and when provided with the opportunity, language testers are often just one voice among many. Advice may easily be disregarded in favour of other advice, or policies used by other, similar, agencies or organisations, as we have seen above.

The study has implications for the training of language testers. There are a number of different possibilities of improving the training that is currently offered, to prepare young language testers better for situations in which they are asked to provide expert advice. I have created the following list of ideas:

- The creation of a repository of case studies of similar work to that described in the three cases above, and that this is made accessible to young language testers; comments on the limitations encountered may also trigger new research projects in the future. Note that the International Language Testing Association (ILTA) recently created the ILTA Advocacy/Public Engagement Award with the intention of recognising and providing models for policy intervention or advocacy that others could follow.
- Where possible, experienced language testers could invite younger colleagues and students to be part of larger project teams that provide policy-advice, either

in a shadowing capacity or more directly associated with the work that is required.

- Roundtable discussions at conferences could be convened to gather a group of experienced language testers to share similar policy-advice situations and evaluate these along similar lines to those in this article.
- The training of language testers should include a range of scenarios that are modelled on real-life situations and that require groups of students to discuss what they would do if they were invited into the policy-making context. Such scenario-based teaching, or problem-driven enquiry would lend itself well to discussions around how to become more policy-responsible language testers.
- Training of language testers should prepare graduates for the fact that advice may not always be heeded due to factors beyond their control. Language testers need to understand that being professional involves the provision of the best possible service within the limits of one's role (see also final annotation bullet-point in ILTA Code of Ethics above).
- Training of language testers may also focus on how to position oneself to give advice, given that there may be missed opportunities to contribute when policy windows briefly open for such possibilities.

The list above is likely to be incomplete and presents just some ideas among many possible others. What most of these items have in common is that they attempt to improve the policy literacy of language testers and applied linguists (Deygers et al., 2021; Pill & Harding, 2013).

Conclusion

By drawing on three recent policy-advice cases, I have attempted to delve into the issues associated with providing expert advice in more detail. I chose to focus on four aspects in each case, but I could have chosen others, such as the time available to provide the advice, how much room there was for negotiation around the advice that was required, or what form our advice took (detailed report, verbal response, etc). The discussion shows that the reach of our advice was not always easy to establish and may in most cases have been limited. A better understanding of why advice was not taken up and why the impact of this work may have been limited in certain cases is an area that needs to be further explored and would help the field to progress in terms of policy literacy and policy impact. The reasons for limited impact in certain contexts may vary due to practical concerns, political compromises, voices of more powerful lobby groups, budget, impact on public opinion, or other competing policies making changes difficult. I have

argued in this paper that young language testers often have limited opportunities for training for such situations, and that further work in this area should be done to prepare graduates better for the situations they may encounter in the future, and for the limited impact their work may have.

References

- Bacchi, C. (2009) *Anaysing policy: What's the problem represented to be?* Sydney: Pearson Education.
- Berthele, R. (2019). Policy recommendations for language learning: Linguists' contributions between scholarly debates and pseudoscience. *Journal of the European Second Language Association*, 3(1), 1-11. <https://doi.org/10.22599/jesla.50>
- Brown J. D., Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing* 25(3):349-383. <https://doi.org/10.1177/0265532208090157>
- Chapelle, C. (2013). Introduction. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley-Blackwell.
- Deygers, B., Bigelow, M., Lo Bianco, J., Nadarajan, D., & Tani, M. (2021, online first). Low Print Literacy and Its Representation in Research and Policy. *Language Assessment Quarterly*. <https://doi.org/10.1080/15434303.2021.1903471>
- Deygers, B., & Malone, M. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, 36(6), 347-368. <https://doi.org/10.1177/0265532219826390>
- Elder, C., Knoch, U., & Harradine, O. (2019). Language requirements for Australian citizenship: Insights from a senate enquiry. In C. Roever & G. Wigglesworth (Eds.), *Social perspectives on language testing*. Peter Lang.
- Fischer, F., & Gottweis, H. (2013). The argumentative turn in public policy revisited: Twenty years later. *Critical Policy Studies*, 7(4), 425-433. <https://doi.org/10.1080/19460171.2013.851164>
- Furedi, F. (2011). Introduction. In B. Molesworth, R. Scullion & E. Nixon (Eds.), *The marketization of higher education and the student as consumer*. Routledge.
- Gan, Z., Stapleton, P., & Yang, C. C. R. (2015). What happens to students' English after one year of English-medium course study at university? *Applied Language Learning*, 25(1&2), 71-91.
- Grabe, W. (2010). Applied Linguistics: A twenty-first-century discipline. In R. B. Kaplan (Ed.), *The Oxford Handbook of Applied Linguistics* (2nd ed.). Oxford University Press.
- Humphreys, P., Haugh, M., Fenton-Smith, B., Lobo, A., Michael, R., & Walkinshaw, I. (2012). *Tracking international students' English proficiency over the first semester of*

- undergraduate study. *IELTS Research Reports Online Series 2012/1*. IELTS Australia Pty Limited, Canberra.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: a focus on language assessment courses. *Language Testing*, 25(3), 385-402. <https://doi.org/10.1177/0265532208090158>
- International Language Testing Association. (2000; 2018). *Code of Ethics*: International Language Testing Association.
- Jann, W., & Wegrich, K. (2007). Theories of the policy cycle. In F. Fischer, G. J. Miller & M. S. Sidney (Eds.), *Public Policy Analysis: Theory, politics and methos*. Boca Raton, London, New York: CRC Press Taylor and Francis Group. <https://doi.org/10.1201/9781420017007.pt2>
- Jin Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing* 27(4):555-584. <https://doi.org/10.1177/0265532209351431>
- Kim, C., Wilcher, R., Petruney, T., Krueger, K., Wynne, L., & Zan, T. (2018). A research utilisation framework for informing global health and development policies and programmes. *Health Research Policy and Systems*, 16(1), 9. <https://doi.org/10.1186/s12961-018-0284-2>
- Knoch, U., McNamara, T., & Elder, C. (2017). *Submission to the Senate enquiry on the Australian Citizenship Legislation Amendment (Strengthening the requirement for Australian citizenship and other measures) Bill 2017*. Language Testing Research Centre: University of Melbourne.
- Knoch, U., Rouhshad, A., Oon, S. P., & Storch, N. (2015). What happens to ESL students' writing after three years of study at an English medium university? *Journal of Second Language Writing*, 28, 39-52. <https://doi.org/10.1016/j.jslw.2015.02.005>
- Knoch, U., Rouhshad, A., & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing*, 21(1), 1-17. <https://doi.org/10.1016/j.asw.2014.01.001>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100-120. <https://doi.org/10.1080/15434303.2019.1674855>
- Lo Bianco, J. (2019). Talking to the pollies: Academic researchers and public officials. In C. Roever & G. Wigglesworth (Eds.), *Social perspectives on language testing*. Peter Lang.

- Macqueen, S., & Ryan, K. (2019). Test mandate discourse: debating the role of language tests. In C. Roever & G. Wigglesworth (Eds.), *Social perspectives on language testing*. Peter Lang.
- O'Loughlin, K., & Arkoudis, S. (2009). *Investigating IELTS exit score gains in higher education*. IELTS Research Reports Vol. 2.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402. <https://doi.org/10.1177/0265532213480337>
- Storch, N., & Hill, K. (2008). What happens to international students' English after one semester at university? *Australian Review of Applied Linguistics*, 31(1), 04.01-04.17. <https://doi.org/10.2104/aral0804>
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36. <https://doi.org/10.1017/S0267190509090035>
- Wray, A., & Wallace, M. (2015). Developing research expertise in applied linguistics. *International Journal of Applied Linguistics*, 166(1), 3-36. <https://doi.org/10.1075/itl.166.1.01wra>