

Developing and validating language proficiency standards for non-native English speaking health professionals

Catherine Elder¹, Tim McNamara¹, Robyn Woodward-Kron², Elizabeth Manias³,
Geoff McColl², Gillian Webb³, John Pill⁴, Sally O'Hagan⁴

¹School of Languages and Linguistics, University of Melbourne, Australia

²Melbourne Medical School, University of Melbourne, Australia

³School of Health Sciences, University of Melbourne, Australia

⁴Language Testing Research Centre, University of Melbourne, Australia

Background

English proficiency is very important for effective healthcare communication. This importance is reflected in the requirement that all overseas trained non-native health professionals seeking to practise in Australia should meet stipulated minimum standards on a recognised test of English proficiency. Nevertheless, the limited communicative skills of some health professionals who have met the English language requirement have been identified as a potential obstacle to effective communication and, by implication, to the delivery of high standards of care (Birrell & Schwartz, 2006; Hawthorne & Birrell, 2002). Overseas trained health professionals play a crucial role in meeting Australia's workforce shortages (Australian Government Department of Health and Ageing, 2008; Barton, Hawthorne, Singh & Little, 2003; Douglas, 2008); therefore, these language issues need to be systematically investigated.

Such issues raise the question of whether the criteria used to assess performance on the currently recognised tests of English are sufficiently aligned to what matters for effective performance in the workplace and whether the passing standards required in terms of these criteria are appropriate. A three-year study funded by the Australian Research Council and the Occupational English Test Centre (LP0991153) explored this question as it applied to the Occupational English Test (OET) (McNamara, 1996). The OET is an English language screening test for overseas-trained health professionals applying to practise in Australia, New Zealand and, more recently, Singapore. Achieving the mandated standard on the OET is regarded by twelve health professions as evidence of adequate English proficiency for healthcare purposes. The test is

occupation-specific in the sense that it is designed to replicate critical communication tasks between health professionals and patients. The speaking test, for example, which is the focus of this study, involves role-plays built around clinical scenarios in which a health professional (the test candidate) interacts with a patient (the test interlocutor). However, the criteria currently used to assess speaking performance – namely FLUENCY, INTELLIGIBILITY, RESOURCES OF GRAMMAR AND EXPRESSION, and APPROPRIATENESS OF LANGUAGE – are linguistic in nature, and as such do not necessarily reflect what is deemed important by health professionals for effective clinical communication. While a further OVERALL COMMUNICATIVE EFFECTIVENESS criterion allows for a holistic assessment of candidates' spoken performance, this criterion is defined generically rather than in health-related terms.

The study

The study was carried out by a multidisciplinary team comprising applied linguists and health professional experts from three of the twelve health professions served by the OET (Medicine, Nursing and Physiotherapy). It involved three research phases: identifying indigenous criteria, applying the indigenous criteria to the assessment of performance on the OET speaking test, and setting new standards against these criteria.

The first phase aimed to identify health professionals' view of spoken communication drawing on the notion of "indigenous assessment" (Jacoby, 1998; Jacoby & McNamara, 1999). It did so by eliciting judgements from health professionals about the qualities of effective and less effective communication displayed in interaction between health professionals and patients in both simulated and workplace training contexts. An analysis of the health professionals' commentary on these interactions revealed a set of recurring themes, as described in Elder et al. (2012), Woodward-Kron et al. (2012) and O'Hagan et al. (forthcoming). These themes formed the basis for a conceptual model representing the aspects of a health professional's performance in the health professional-patient consultation that are valued by expert practitioners (Pill, forthcoming).

In the second phase, these valued aspects of communication were considered in relation to the current criteria used to assess performance on the OET speaking test role-plays. While the current linguistic criteria did in fact feature among the qualities valued by health professionals, other professionally relevant features mentioned in their feedback were not well represented on the OET. The feasibility of adding two new criteria, CLINICIAN ENGAGEMENT and

MANAGEMENT OF INTERACTION, based on the health professionals' feedback, was explored using a sample of speaking test audio-recordings from the OET assessment bank. The two new criteria were introduced in place of the holistic OVERALL COMMUNICATIVE EFFECTIVENESS criterion currently used to assess task performance. A group of active OET assessors were trained in the application of these new criteria alongside the existing linguistic ones, using a checklist of indicators developed expressly for the project and drawing on insights from the data gathered in its first phase. Analyses showed that OET assessors were able to understand and apply the new criteria successfully and that combining these criteria with the existing ones yielded reliable measurements of candidate performance.

The third and final phase of the study was geared to setting appropriate passing standards on the OET. Health professionals from Medicine, Nursing and Physiotherapy took part in discipline-specific standard-setting workshops where they were issued with pre-scored audio-recorded OET speech samples and asked to classify them according to their perceptions of each candidate's readiness to cope with the speaking demands of their healthcare setting. The collective classifications of these health professionals were used to establish revised cut-scores on the OET. The estimated impact of the proposed new cut-scores varied somewhat across professions, yielding lower pass rates in some cases, but also resulting in more candidates being classified as highly competent.

The study has practical implications for the OET. It seems appropriate to consider changes to the OET criteria and passing standards, based on the research outcomes reported above, although further work is needed to ascertain that the same criteria and standards can be applied to the health professions not included in this study. If the new criteria are adopted, consideration needs to be given to how performance is reported – whether as two scores representing linguistic competence and professionally-oriented communicative competence respectively or as a single combined score. Any decisions must, however, wait on deliberations within the OET Centre and on consultation with relevant stakeholders.

The study's contribution to the theory and practice of LSP (language for specific purposes) testing more generally is significant and will be elaborated in presentations and publications stemming from the project. The approach to identifying 'indigenous criteria', by eliciting and codifying feedback from occupational experts on a range of workplace-related interactions could serve as a model for occupations in fields other than healthcare, where specific-purpose tests are adopted. The process of translating what matters to health professional

experts into criteria that can be applied by language experts to assess performance on a language test is highly original and again holds promise for a field which has thus far tended to define language proficiency in narrowly linguistic terms. Finally, the performance-based standard-setting procedures adopted in this study, while not original in themselves (e.g., see previous accounts in Cizek & Bunch, 2007; Cizek, Bunch & Koons, 2004), have thus far been neglected in published reports (and this is particularly true for performance-based methods of assessment such as the role-play scenarios used in the OET). Since decisions about who is or is not fit to practise in healthcare settings have serious consequences for those affected (in this case, the practitioners themselves and their patients and colleagues), documentation of the steps taken to ensure that these decisions are legally and academically defensible is of paramount importance and should be a requirement for all high-stakes tests.

Acknowledgements

The research team is grateful for the contributions of staff at the OET Centre and of the group of OET assessors. We also wish to thank the health professionals who participated in the project workshops. We are indebted to the patients who consented to the recording of their interactions with health professionals and to the students and supervisors who allowed us to record their feedback sessions. The project advisory team provided valuable feedback and recommendations, and colleagues at the Language Testing Research Centre were always generous with their intellectual input and support.

References

- Australian Government Department of Health and Ageing. (2008, April). *Report on the audit of health workforce in rural and regional Australia*. Canberra: Commonwealth of Australia.
- Barton, D., Hawthorne, L., Singh, B. & Little, J. (2003). Victoria's dependence on overseas-trained doctors in psychiatry. *People and Place*, 11(1), 54–64.
- Birrell, B. & Schwartz, A. (2006). Accreditation of overseas trained doctors: The continuing crisis. *People and Place*, 14(3), 1–11.
- Cizek, G. J. & Bunch, M. B. (2007). *Standard setting: A guide to establish and evaluating performance standards on tests*. London: Sage Publications.
- Cizek, G. J., Bunch, M. B. & Koons, H. (2004). Setting performance standards: Contemporary methods. *Educational Measurement*, 23(4), 31–49.

- Douglas, S. (2008). The registration and accreditation of international medical graduates in Australia – a broken system or a work in progress? *People and Place*, 16(2), 28–40.
- Elder, C., Pill, J., Woodward-Kron, R., McNamara, T., Manias, E., Webb, G. & McColl, G. (2012). Health professionals' views of communication: Implications for assessing performance on a health-specific English language test. *TESOL Quarterly*, 6(2), 409–419.
- Hawthorne, L. & Birrell, B. (2002). Doctor shortages and their impact on the quality of medical care in Australia. *People and Place*, 10(3), 55–68.
- Jacoby, S. W. (1998). *Science as performance: Socializing scientific discourse through conference talk rehearsals*. Unpublished PhD thesis. University of California, Los Angeles.
- Jacoby, S. & McNamara, T. (1999). Locating competence. *English for Specific Purposes*, 18(3), 213–241.
- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- O'Hagan, S., Manias, E., Elder, C., Pill, J., Woodward-Kron, R., McNamara, T., McColl, G. & Webb, G. (forthcoming). What counts as effective communication in nursing? Evidence from nurse educators' and clinicians' feedback on nurse interactions with simulated patients. *Journal of Advanced Nursing*.
- Pill, T. J. H. (forthcoming). *What doctors value in consultations and the implications for specific-purpose language testing*. PhD thesis, University of Melbourne.
- Woodward-Kron, R., van Die, D., Webb, G., Pill, J., Elder, C., McNamara, T., Manias, E. & McColl, G. (2012). Perspectives from physiotherapy supervisors on student–patient communication. *International Journal of Medical Education*, 3, 166–174.