

1

Introduction

Beverly Baker

University of Ottawa, Canada

Lynda Taylor and I are proud to introduce Studies in Language Testing Volume 56, *Language Assessment Literacy and Competence: Case Studies from Around the World*. This volume is a companion to Volume 55, *Research and Reflections from the Field*. Together the two volumes focus on current understandings and practice in the development of language assessment literacy (LAL) and competence as reflected in a series of empirical research papers, case study reports and individual reflections.

Here I describe how this volume came about, sharing our understandings of case studies together with our motivation in presenting a collection of cases to complement the traditional chapters of Volume 55. A part of this motivation includes redirecting attention to the centrality of *context* in our field, something that case studies are especially effective in highlighting and exploring. With this volume we aimed to dispel narrow perceptions of the usefulness of case studies, drawing the reader's attention to some of the key contributions of this particular collection of cases to scholarship in language assessment literacy and competence.

In our planning we did not originally set out to dedicate a volume to case studies *per se*. In mid-2021, while reading through the dozens of submissions for our proposed volume, Lynda and I were struck by the diversity represented in the abstracts in terms of geography, data sources, collaborators, and approaches to investigating language assessment knowledge building. We saw the potential to profile some of these scholars and their work in a new way. As a result, we contacted a selection of these scholars to ask if they would consider developing their submission not as a traditional book chapter but as a report of their particular case – shorter but with more concentrated details of their specific context than they might normally report. While many of our authors did not necessarily set out to do case study research, this does not need to be done *a priori*: as Yin (2013) reminds us, cases can be conceptualised as a way to describe situated phenomena. Our contributors were willing to reconceptualise their work in this way, providing the particularities of their situation and endeavoring to share what might resonate for others in their own circumstances across our international community. We commend them for taking up this challenge. Their successful efforts here can encourage others to feel confident about doing something similar in their own contexts.

Language Assessment Literacy and Competence Volume 2

Defining and operationalising case studies

In the words of Gerring, ‘Regretfully, the term “case study” is a definitional morass’ (2004:342), suggesting that it is so amorphous a concept as to defy description. Despite this, he provides what we find to be a very useful definition, straightforward yet broad in scope: ‘A case study is best defined as an in-depth study of a single unit (a relatively bounded phenomenon) where the scholar’s aim is to elucidate features of a larger class of similar phenomena’ (2004:341).

This idea of relative as opposed to absolute boundaries allows for remarkable conceptual flexibility, and this volume effectively demonstrates the various forms a case can take, such as the following:

- the case of *a single school* that implemented an innovative approach to FL learning in Australia, where the teachers and students are learning the language at the same time (**Kathryn Macfarlane**, Chapter 6)
- *an organisation* (ALTE) with a mission of disseminating knowledge and expertise in language assessment to a wide variety of stakeholders, through the provision of resources, educational initiatives and policy collaboration and advocacy (**Graham Seed, Waldemar Martyniuk and Lorenzo Rocca**, Chapter 10)
- *an activity* to design analytic scales for teachers in various schools within a public language school network in the Canary Islands (**Joaquín M Cruz Trapero**, Chapter 11)
- *a group of three teacher trainers* on the Cambridge CELTA (a teaching certification course), using an innovative ‘lesson study’ approach for their professional development (**Magnus Coney and Ben Naismith**, Chapter 12).

Many of the cases reported in the volume concern groups of teachers, at various levels and stages of their careers, spanning the globe. Such cases represent *local testing and assessment* defined by Dimova, Yan and Ginther as ‘one whose development is designed to represent the values and priorities within a local instructional program and designed to address problems that emerge out of a need within the local context in which the test will be used’ (2020:1). For example, **Frank Giraldo** (Chapter 7) investigated a group of practising high school teachers in one area of Colombia, tailoring training to the teachers’ specific needs and goals. **Valia Spiliotopoulos, Saskia Van Viegen, Margaret Early and Connie Lam James** (Chapter 2) share the activities of a university–school district professional learning community working on the implementation of new English language learning standards in British Columbia, Canada. **Olga Ukrayinska** (Chapter 3) explored the development of self-assessment skills of pre-service English and French teachers in Ukraine. We applaud her for succeeding in contributing to this

Introduction

volume despite a number of practical obstacles. We appreciate her continued dedication and that of her colleagues to teacher education in the face of ongoing conflict.

The cases presented here are not all delineated by geography. The projects described by **Carolyn Westbrook** and **Richard Spiby** (Chapter 9) as well as by **Seed et al** (Chapter 10) and **Rob Playfair** and **Eddie Cowling** (Chapter 14) do not focus on a specific geographical context. Instead, they report initiatives designed for groups with a specific common need. Contexts here can also be described in terms of ‘micro-contexts at the level of the individual or macro-contexts of the wider national, political, and educational system in which that individual is located’ (Lai and Fjørtoft 2022:1). In our volume, **Jiyoan Lee** (Chapter 5) investigates the reflections of pre-service teachers with a classroom-based project in the United States (a micro context) while **Dina Vilcu**, **Koen Van Gorp** and **Marta Garcia** (Chapter 13) explore the revision of two English for Specific Purposes tests in Romania and Spain with reference to an international guidelines document (perhaps best described as a meso context). Regarding large-scale, commercially available standardised tests (macro contexts), **Mojtaba Heydari** and **Fahimeh Marefat** (Chapter 4) employed an innovative data mining procedure to explore comments from a social networking discussion group of test-takers in order to uncover their (mis)conceptions of IELTS as well as their concerns with the test taking experience – as expressed to their peers rather than to researchers.

The importance of case studies in advancing knowledge

The definitions of cases above refer not just to how cases can be delineated, but to their uses and their usefulness in contributing to knowledge. Stake (1995) defines cases as *intrinsic* (studying a unique instance), *instrumental* (indicative of larger phenomena) and *collective* (used as a collection to paint a portrait of an issue). We see these categories as complementary rather than mutually exclusive. Firstly, the *intrinsic* nature of cases means priority is given to the insights that can be gained from studying a relatively limited object or grouping *in situ*. Flyvbjerg writes: ‘The advantage of a case study is that it can “close in” on real-life situations and test views directly in relation to phenomena as they unfold in practice’ (2006:235). Crowe et al echo this sentiment, calling case study research ‘... an in-depth appreciation of an issue, event or phenomenon of interest, in its natural real-life context’ (2011:1).

Cases are *instrumental* in that they ‘illustrate broader lessons that may be learnt’ (Crowe et al 2011:1), and *collective*, in the sense that knowledge of collections of cases can be viewed as the basis for the development of expertise. This is similar to what Yin (2013) terms ‘analytic generalization’,

Language Assessment Literacy and Competence Volume 2

a process of abstracting from a case outward. Flyvbjerg argues that ‘... a scientific discipline without a large number of thoroughly executed case studies is a discipline without systematic production of exemplars, and a discipline without exemplars is an ineffective one’ (2006:219). Stating that the case study ‘may be central to scientific development’ (2006:228), he provides examples of research that has advanced humankind through individual experiments and case studies as opposed to formal generalisation and hypothetical-deductive models: ‘That knowledge cannot be formally generalized does not mean it cannot enter into the collective process of knowledge accumulation in a given field or in a society. A purely descriptive, phenomenological case study ... can certainly be of value in this process and has often helped to cut a path toward scientific innovation’ (2006:227). He takes the example of the law of gravity, which was established not through controlled experimentation with random samples, but through the selection of the two cases of lead and a feather to determine whether they would fall with equal velocity.

As can be inferred above, one’s understanding of the term ‘generalisable’ depends on one’s epistemological positioning as critical, interpretive or positivist (Crowe et al 2011). Flyvbjerg (2006) describes a positivist orientation as one which privileges ‘context-independent’ knowledge. An extreme view of this would be the concept of an ideographic universe, described by Gerring where ‘absolutely nothing can be learned about one unit by studying another’ (2004:351). Those taking critical and interpretivist positions tend to view knowledge as essentially indivisible from the context in which it is generated, a view that is consistent with the current post-positivist orientation of our field and in the social sciences more generally. Judging from the recent proliferation of context-bound empirical studies in language assessment, we make two observations: i) that scholars in our field are taking a generally interpretative approach, acknowledging the inseparability of context from their discoveries, and ii) that despite this, they do see the instrumental and collective value in sharing these context-bound experiences with others.

In other words, we see evidence from the very existence of this work that scholars across our field adhere to the belief that ‘a context-specific problem ... can be solved while advancing methodological and theoretical knowledge in a range of domains ...’ (Lai and Fjørtoft 2022:2). Of course, discoveries in one context may require a certain amount of imagination to apply to other settings. Lai and Fjørtoft describe the concept of ‘boundary crossing’ from educational theory as the highly productive movement of innovations among contexts: ‘ideas or practices arising in one educational context are therefore not seamlessly integrated into others; instead, they are subject to processes of reinterpretation, hybridization, or transformation’ (2022:2).

This discussion of the value and potential of cases brings us to a further observation regarding an unfortunate convention we still see regularly

Introduction

in our field: after scholars have presented their in-depth case study at a conference or in a publication, and have argued, often convincingly, for the insights their discoveries can offer their peers and the knowledge base of their field, they then finish by stating that their case study approach may be considered a ‘limitation’ of their research – that because it is a case, it ‘cannot be generalized’. For us, this type of statement amounts to conducting a study through a critical or interpretive framework and then veering into positivism at the last minute, at the very moment the researcher is arguing for the value of their work! The intrinsic, instrumental and collective benefits of case studies which motivated the study in the first place are discounted with reference to some sort of context-free standard. Such statements risk perpetuating the idea that the gold standard in research of any kind is the randomised controlled experimental study, and that knowledge is only generated by conceiving of a case as a representative sample which can represent the population as a whole. Can we rid ourselves of this type of self-defeating and incoherent statement? Flyvbjerg (2006) challenges this misunderstanding of generalisation and its role in knowledge production, contending the following:

Social science has not succeeded in producing general, context-independent theory and, thus, has in the final instance nothing else to offer than concrete context-dependent knowledge ... Predictive theories and universals cannot be found in the study of human affairs. Concrete, context-dependent knowledge is, therefore, more valuable than the vain search for predictive theories and universals ... formal generalization is overvalued as a source of scientific development, whereas the “force of example” is underestimated (223–228).

He evokes Beveridge, who contends that there ‘are more discoveries stemming from the type of intense observation made possible by the case study than from statistics applied to large groups’ (1951:236).

Insights gained from this collection

In summary, this collection demonstrates strong motivation to learn more about language assessment, at least for the diverse actors involved in language education. Desire to learn resonates through multiple contributions here. In terms of the construct, we see the term ‘literacy’ still used most often, with other terms such as ‘competence’ occasionally applied. Regardless of the terms used, it does not appear to be a controversial idea that different stakeholders’ knowledge can be described in terms of differing but complementary profiles. We also observe a movement from questions related to the constituent elements of LAL – the WHAT (e.g., Kremmel and Harding 2020) – to questions of HOW to facilitate this development.

Language Assessment Literacy and Competence Volume 2

We see an awareness that there is no ‘one size fits all approach’ to LAL development for all stakeholders. In this volume, LAL development is addressed with a nascent professional learning community (**Spiliotopoulos et al**, Chapter 2), through Learning Study (**Coney and Naismith**, Chapter 12), and with online interactive workshops (**Giraldo**, Chapter 7), to name just some examples. We also see examples of development through collaboration, such as between teachers and their learners (**Lee**, Chapter 5, **Macfarlane**, Chapter 6) and among testing experts, teachers, and subject matter experts in the development of a specific purposes medical language assessment (**Vilcu et al**, Chapter 13). In addition, the voices of test-takers themselves are recognised as sources of learning for language test developers themselves (seen in **Heydari** and **Marefat**, Chapter 4, among others).

Self-reports seem to be the preferred investigative tool represented here. For example, in their efforts to build an international piloting network amongst English for Academic Purposes practitioners, **Playfair and Cowling** (Chapter 14) used interviews to learn about the needs and challenges of their colleagues. In addition to self-reports, we see a place for an increase of observational studies including observation of formative assessment practices in the classroom (such as **Susan Sheehan** and **Vivien Berry**, Chapter 8, and **Ukrayinska**, Chapter 3) and collection of assessment-related products (see **Cruz Trapero**’s work in Chapter 11, where rating scales created by teachers serve as proof of learning).

These represent only a few of our own insights about the collection as a whole and what it represents for us and our own understandings, as well as for the field of language assessment more broadly. However, in the interests of ‘boundary crossing’, we invite our readers to interpret and draw their own insights from this collection according to their individual needs and interests, perhaps even following up with our authors for any details that might be helpful for a productive reimagining of this work in other contexts. Above all, we believe this collection represents a recognition and affirmation of the value of cases in our field, both as a scholarly language and as a way of enabling our progress as a collective.

References

- Beveridge, W I B (1951) *The Art of Scientific Investigation*, London: Heinemann.
- Crowe, S, Cresswell, K, Robertson, A, Huby, G, Avery, A and Sheikh, A (2011) The case study approach, *BMC Medical Research Methodology* 11, Article 100.
- Dimova, S, Yan, X and Ginther, A (2020) *Local Language Testing: Design, Implementation, and Development*, London: Routledge.
- Flyvbjerg, B (2006) Five misunderstandings about case-study research, *Qualitative Inquiry* 12 (2), 219–245.

Introduction

- Flyvbjerg, B (2021, May) Yes, you can generalize from a case study, despite what your supervisor says, *Geek Culture*, available online: medium.com/geekculture/yes-you-can-generalize-from-a-case-study-f8000bd647b9
- Gerring, J (2004) What is a case study and what is it good for?, *The American Political Science Review* 98 (2), 341–354.
- Kremmel, B and 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.
- Lai, M K and Fjørtoft, H (2022) Problematising and understanding contexts in educational research: An international perspective, *Studies in Educational Evaluation* 74, 1–4.
- Stake, R (1995) *The Art of Case Study Research*, London: Sage Publications.
- Yin, R (2013) *Case Study Research: Design and Methods* (Fifth edition), Thousand Oaks: Sage Publications.