

RESEARCH TIMELINE

## Classroom-based assessment

Jo Lewkowicz<sup>1\*</sup> and Constant Leung<sup>2</sup>

<sup>1</sup>King's College London, UK/University of Warsaw, Poland and <sup>2</sup>King's College London, UK

\*Corresponding author. Email: [jo.lewkowicz@kcl.ac.uk](mailto:jo.lewkowicz@kcl.ac.uk)

### Introduction

The assessment of students in the classroom has been going on since time immemorial. What is comparatively recent, however, is the systematic study of classroom-based assessment (CBA). The term 'CBA' has been putatively linked to Michael Scriven's (1967) work on formative and summative evaluation. However, current interest in such assessment and how it is enacted has, to a large extent, been prompted by shifts in educational policy in various contexts and evolving education systems. This, in turn, has led to the increase in research activity that is detailed in the timeline that follows. At the same time, considerable effort has been exerted by various governments and professional associations into the development of CBA frameworks, but as publications related to these are not strictly research documents, a separate list of examples is provided as supplementary material.

Initially, research interest in the nature and enactment of formative CBA came from the field of general education. For example, in the late 1980s, as part of a radical overhaul of the school education system in England that placed emphasis on central government control of the curriculum and of assessment, the Task Group on Assessment and Testing (TGAT) drew up a report that reflected an attempt to, *inter alia*, reconcile the conflicting demands of high-stakes public reporting of student performance in accordance with statutory curriculum specifications, and educationally-oriented assessment that reports student progression in the light of teaching (and therefore learning) experience (TGAT, 1987). Paragraph 5 of the report stated that the results of assessment 'should provide a basis for decisions about pupils' further learning needs: they should be formative'.

The chair of the TGAT task group, Paul Black, later published a seminal paper on formative assessment with Dylan Wiliam (Black & Wiliam, 1998). This seminal 'state of play' review of research on classroom-based formative assessment covered 681 publications which reported on a range of studies with diverse research designs and methodologies investigating formative assessment schemes and practices at different educational levels and settings internationally. Some of the studies had experimental and control groups, others were classroom-based and teacher-led. The analysis and commentary in the Black and Wiliam paper foregrounded issues such as quality of teacher–student interaction, feedback, the role of the student in assessment, and self and peer assessment. Many of these issues resonate with the current work in formative assessment. The authors' list of issues to be taken into account by researchers in investigating formative assessment is as relevant today as it was then. Many of these have been taken up by those researching CBA within the field of languages education with some reflected in the present timeline.

Educational reform in England in the 1980s provided the impetus for investigations into how teachers assess in situations where English is used as an additional language particularly at the primary and secondary school levels of education. But it was not only England that towards the turn of the century was experiencing educational reform: notable examples of reform include the introduction of the National Certificate of Educational Achievement in New Zealand in 2002 (see East & Scott, 2011, for a historical account), the introduction of school-based assessment (SBA) as an element of secondary school assessment in Hong Kong from 2001, as well as the introduction of Scottish

National Standardised Assessments since 2017. What all these initiatives have in common is a commitment to the assessment for learning (AfL) as well as a growing interest in understanding how such assessment is implemented at the classroom level. This, in turn, has exposed existing tension between national policy recommendations and challenges of classroom implementation. It has also drawn attention to the need for improved assessment literacy for teachers, many of whom have continued to rely heavily on summative assessment practices despite the fact that limitations of summative assessment had been unequivocally demonstrated. Wiliam (2001), for example, was able to show that from a mathematical perspective the results of high-stakes tests may produce highly unreliable results for individual students even when they are deemed reliable for the group as a whole.

The formative–summative assessment dichotomy was for some time portrayed over-simplistically. Case study research of individual classrooms has shed light on the complexities of formative CBA in general and more specifically in language education which has to work with a variety of pedagogic concepts and learning theories. This kind of research has highlighted the limited applicability of the psychometric paradigm dominant within summative, standardised assessment to account for unplanned, often context and content-embedded implicit assessment that is part and parcel of CBA. Such research has provided a basis for the development of CBA frameworks that take account of teacher and student interactions. Unlike in psychometrically oriented assessments, CBA is predominantly located within the classroom, thus making it important to take account of participants' perspectives and understandings in relation to curriculum and assessment requirements.

CBA has developed along two main trajectories: theory building to better understand its conceptual basis (construct) and analysing on a moment-by-moment basis how it is enacted in classrooms at different educational levels and in different world locations. These, therefore, form the main themes of our timeline. In this timeline our working definition of CBA is: Any teacher-led classroom activity designed to find out about students' performance on curriculum tasks that would yield information regarding their understanding as well as their need for further support and scaffolding with reference to their situated learning needs. (We recognise that not all CBA is teacher-led and that self and peer assessment may play an important role, but these aspects of CBA would warrant a timeline in their own right.) We further recognise that there have been shifts in the use of terms such as 'foreign language', 'second language' and 'additional language' in recent times. We will use 'additional/second language' as a catch-all term, but we will use 'foreign' and 'second' where it is representationally important to signal historical accuracy. Our timeline contains works from additional/second language research, and from the field of education more generally where appropriate.

Thus, our selection of themes is as follows:

- A understanding of the conceptual basis (construct) of CBA
- B implementation and enactment of CBA

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/S0261444820000506>

## References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- East, M., & Scott, A. (2011). Assessing the foreign language proficiency of high school students in New Zealand: From the traditional to the innovative. *Language Assessment Quarterly*, 8(2), 179–189.
- Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Curriculum evaluation* (pp. 39–83). Chicago, IL: Rand McNally. American Educational Research Association. (monograph series on evaluation, no. 1).
- TGAT (Task Group on Assessment and Testing). (1987). *National Curriculum: A report*. London Department of Education and Science and the Welsh Office.
- Wiliam, D. (2001). Reliability, validity and all that jazz. *Education 3–13: International Journal of Primary, Elementary and Early Years Education*, 29(3), 17–21.

**Jo Lewkowicz** is a visiting lecturer at King's College London. She is also a consultant in language testing at the University of Warsaw in Poland where she advises on the implementation of high-stakes language examinations that are part of the

graduation requirement for all students at the university. She has previously taught at tertiary level in a range of countries including Egypt, Kenya, mainland China, Hong Kong and Armenia. Her research interests include language testing and assessment with a particular focus on the authenticity of assessment as well as the development of academic literacy within and across university disciplines.

**Constant Leung** is Professor of Educational Linguistics in the School of Education, Communication and Society, King's College London. Before taking up teaching positions in higher education he taught in schools and worked as an advisory teacher and manager in local government. He was the founding chair of the National Association for Language Development in the Curriculum. His research interests include education in ethnically and linguistically diverse societies, additional/second language curriculum and assessment, language policy and teacher professional development. He serves as Co-Editor of *Language Assessment Quarterly*, Editor of Research Issues of *TESOL Quarterly*, and as a member of the editorial boards of *Australian Review of Applied Linguistics*, *Language and Education*, and the *Modern Language Journal*. He is a Fellow of the Academy of Social Sciences (UK). His work in developing the English as an Additional Language Assessment Framework for Schools (funded by the Bell Foundation) won the 2018 British Council ELTons international award for innovation.

Year	References	Annotations	Theme (s)
1998	Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. <i>Assessment in Education</i> , 5(1), 85–102.	<b>Perrenoud</b> , engaging the English language literature from a French research perspective, argues that much scholarly work (up to the time of writing) in formative CBA focussed on evaluative feedback, assuming that feedback would promote learning. Perrenoud argues that such an assumption is not necessarily ‘safe’ as student responses cannot be assumed. While this paper is concerned with educational assessment generally, many of the issues discussed have conceptually foreshadowed a good deal of the more recent critically minded work in the field of additional/second language education.	A
2000	Hasselgren, A. (2000). The assessment of English ability of young learners in Norwegian schools: An innovative approach. <i>Language Testing</i> , 17(2), 261–277.	<b>Hasselgren</b> provides an early example of formative CBA implementation being trialled for use with young learners in the Norwegian context. She provides a detailed description of how the assessment materials and scoring instruments were developed, and reflects on issues raised by REA-DICKINS & GARDENER (2000).*	B
2000	Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. <i>Language Testing</i> , 17(2), 215–243.	<b>Rea-Dickins and Gardner</b> were among the first to focus attention on the construct of formative assessment. On the basis of feedback from teachers, they demonstrated the complex nature of formative assessment and showed that the demarcation between summative and formative assessment was not clear-cut as previously believed.	A
2000	Teasdale, A., & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? <i>Language Testing</i> , 17(2), 163–184.	With the increasing debates on the educational merits and the advocacy for classroom-based formative assessment in the 1990s (also known as alternative assessment at the time), there was a need, also acknowledged by REA-DICKINS & GARDENER (2000), to reconsider the established conceptual and operational assumptions in second/foreign language testing in terms of validity and reliability. In this paper <b>Teasdale and Leung</b> provide an account of the conceptual and theoretical difficulties in assuming that the psychometric assumptions and principles underlying standardised (usually summative) testing can be elided unproblematically into classroom-based formative assessment. The implications of the incongruences for assessment, pedagogy and policy are discussed.	A
2001	Brindley, G. (2001). Outcomes-based assessment in practice: Some examples and emerging insights. <i>Language Testing</i> , 18(4), 393–407.	<b>Brindley</b> , on the basis of examples from school and adult immigrant education in Australia, draws attention to issues that arise in the implementation of outcomes-based assessment (which may be external, teacher-led assessment or a combination of both). Particularly in the case of teacher-constructed assessment, he shows there is a complex	B

\*Note. Authors’ names are shown in small capitals when the study referred to appears in this timeline.

		interplay of factors that determine how the assessment is operationalised and suggests that teachers need to be adequately prepared for conducting CBA. The issues raised resonate with more recent work, e.g. INBAR-LOURIE & DONITSA-SCHMIDT (2009).	
2001	Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. <i>Language Testing</i> , 18(4), 429–462.	<b>Rea-Dickins</b> draws on the growing CBA literature within general education (see BLACK & WILLIAM, 1998, for a comprehensive overview) to investigate classroom-based teacher assessment in a mainstream English elementary school setting with particular reference to students from an English as an additional language (EAL) background. The discussion provides a context-sensitive account of the complex curricular and pedagogic processes and practices when studying classroom-based teacher assessment.	<b>B</b>
2004	Broadfoot, P., & Black, P. (2004). Redefining assessment of L2 development? The first ten years of Assessment in Education. <i>Assessment in Education</i> , 1(1), 7–27.	The 1990s witnessed the emergence of diverse approaches to educational assessment that were sensitive to local contexts and social practices; formative approaches were part of these developments. There was, at the same time, a hardening of policy support for standardised assessment for public accountability. This authoritative retrospective paper provides an informed view on the various ‘movements’ within the field of educational assessment. In the sections on formative assessment <b>Broadfoot and Black</b> highlight the pitfalls in formative practices in predominantly summative assessment-oriented educational environments, indicating the growth points in conceptualising and developing formative assessments. TEASDALE & LEUNG (2000) made many similar observations from the standpoint of additional/second language assessment.	<b>A</b>
2004	Cheng, L., Rogers, T. & Hu, H. (2004). ESL/ESL instructors’ classroom assessment practices: Purposes, methods and procedures. <i>Language Testing</i> , 21(3), 360–389.	Like DAVISON (2004), <b>Cheng et al.</b> explore teacher assessment practices, through self-report at tertiary level across three locations (Beijing, Hong Kong and Canada). The paper also highlights the variability of CBA across the different social contexts in which assessment is administered.	<b>B</b>
2004	Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australia and Hong Kong secondary schools. <i>Language Testing</i> , 21(3), 305–334.	<b>Davison</b> explores the variability of secondary school teacher beliefs, attitudes and values when assessing student work as well as of the extent to which teachers believe their judgements are legitimate and trusted by stakeholders. Results suggest considerable variability across teachers and teaching contexts (i.e. Australia and Hong Kong).	<b>B</b>
2004	Lantolf, J. P., & Poehner, M. (2004). Dynamic assessment of L2 development: Bringing the past into the future. <i>Journal of Applied Linguistics</i> , 1(1), 49–72.	Dynamic assessment (DA) draws on Vygotskian principles of human learning and development. It is often regarded as distinct from formative assessment in the educational assessment literature; however, they have in common a shared interest in promoting learning. In this paper <b>Lantolf and Poehner</b> argue that conventional approaches to assessment tend to	<b>A</b>

(Continued)

(Continued)

Year	References	Annotations	Theme (s)
		be static since they are designed to identify what a learner has achieved in the past. In contrast, DA works with an individual's learning potential to maximise achievement; in that sense DA is future oriented. The discussion of the interventionist and interactionist approaches is directly relevant to CBA in terms of what is to be taught and learnt.	
2004	Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. <i>Language Assessment Quarterly</i> , 1(1), 19–41.	With the rising interest of researchers (e.g. BRINDLEY, 2001; REA-DICKINS, 2001; TEASDALE & LEUNG, 2000) and policy makers in teacher-based formative assessment, <b>Leung</b> argued for a reconceptualisation of how CBA is investigated, showing the importance of accounting for individual teachers' situated practices. Drawing on empirical data, Leung highlighted the diverse and complex thinking that teachers bring to the classroom assessment process.	A & B
2005	Allal, L., & Lopez, L. M. (2005). Formative assessment of learning: A review of publications in French. In Organization of Economic Co-operation and Development, <i>Formative assessment: Improving learning in secondary classrooms</i> (pp. 241–254). Paris, France: OECD Publishing.	This review of CBA literature provides a useful snapshot of the research conducted in France and Francophone regions of Belgium, Canada and Switzerland from the 1980s to the early 2000s. It offers, inter alia, a classification of the different conceptualizations of CBA and the different types of empirical research; of particular interest is the distinction between <i>remédiation</i> (feedback + correction) and <i>régulation</i> (feedback + adaptation). This review can be read as a companion piece to BROADFOOT & BLACK (2004).	B
2005	Carless, D. (2005). Prospects for the implementation of assessment for learning. <i>Assessment in Education</i> , 12(1), 39–54.	Drawing on two examples of 'early adopters' of AfL within the Hong Kong primary school context, focussing on the work of English as a second language (ESL) teachers, <b>Carless</b> discusses factors likely to facilitate or inhibit the implementation of AfL, highlighting that the time needed for successful implementation to take root is often not accorded by those responsible for educational reform. This paper points to the need for system-wide support for CBA.	A
2006	Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: A learner-focused interpretation. <i>International Journal of Applied Linguistics</i> , 16(2), 163–188.	In contrast to her earlier work (REA-DICKINS, 2001) relying on teacher self-report, in this paper <b>Rea-Dickins</b> looks at learners in dialogue with their teachers and peers as they are being assessed in the classroom. She draws on longitudinal data from the EAL primary classroom in the UK to show how different teaching agendas shape assessments variably. She argues that within instruction there should be a balance of summative and formative assessment so as to provide learners with sufficient opportunities to engage in assessment and develop their language and language learning awareness.	B

2006	Yueming J., Eslami, Z. R., & Burlbaw, L. M. (2006). ESL teachers' perceptions and factors influencing their use of classroom-based reading assessment. <i>Bilingual Research Journal</i> , 30(2), 407–430.	The turn of the century saw a marked growth in interest of ESL/EAL teacher perceptions of CBA at schools (cf. DAVISON, 2004; REA-DICKINS, 2001; TEASDALE & LEUNG, 2000). <b>Yeuming et al.</b> 's paper follows this trend reporting on an investigation of teachers' perceptions of CBA reading assessment among elementary and middle school teachers in the US. It suggests that teachers are more favourably disposed to CBA than to state-mandated standardised tests. However, as state-mandated tests and district policies control assessment, they constrain the implementation of innovative approaches to CBA.	B
2007	Leung, C., & Rea-Dickins, P. (2007). Teacher assessment as policy instrument: Contradictions and capacities. <i>Language Assessment Quarterly</i> , 4(1), 6–36.	In this paper <b>Leung and Rea-Dickins</b> explore policy and practice of assessment at a time of curriculum change in the UK. They highlight the rift between policy makers in search of demonstrating rising educational attainment and the realities of the classroom. Echoing REA-DICKINS & GARDNER (2000), the authors argue for the need to distinguish not only between summative and formative classroom assessment but also between assessing English as a first language and EAL.	B
2007	McKay, P., & Brindley, G. (2007). Educational reform and ESL assessment in Australia: New roles and new tensions. <i>Language Assessment Quarterly</i> , 4(1), 69–84.	The variable impact of educational reform on classroom assessment practices (discussed also in YUEMING ET AL., 2006 in relation to the US; LEUNG & DICKINS, 2007, with reference to the UK.) and the tensions it may create between teachers and management is well illustrated by <b>McKay and Brindley</b> with reference to the Australian context. The authors consider how an outcomes-based system of assessment introduced for accountability purposes impacted on school and adult education. In the ESL school sector, the change appeared to have a negative impact as teachers focussed more on the high-stakes standardised assessment and less on the curriculum and individual learner needs (sentiments echoed by INBAR-LOURIE & DONISTA-SCHMIDT, 2009). In the adult sector where there was no standardised final test, reporting on learner outcomes appeared to be variable and not consistent across teachers but teachers remained more learner-centred in their approach.	B
2009	Inbar-Lourie, O., & Donitsa-Schmidt, S. (2009). Exploring classroom assessment practices: The case of teachers of English as a foreign language. <i>Assessment in Education: Principles, Policy &amp; Practice</i> , 16(2), 185–204.	<b>Inbar-Lourie and Donitsa-Schmidt</b> report on a number of internal as well as external factors that impinge on teachers' employment of CBA and the tensions that exist between the highly centralised, top-down Israeli education system that prizes high-stakes testing and more recent attempts to introduce alternative assessment within the language classroom. Their observations resonate with those of BRINDLEY (2001).	B
2009	Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. <i>TESOL Quarterly</i> , 43(3), 471–491.	<b>Poehner</b> discusses the implementation of group DA and shows how individual learners working in groups may be supported through DA in the classroom. This discussion extends the conceptual frame of formative CBA (see LANTOLF & POEHNER, 2004) that tends to focus on one-to-one teacher–student interaction.	A

(Continued)

(Continued)

Year	References	Annotations	Theme (s)
2011	Cheng, L., Andrews, S., & Yu, Y. (2011). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. <i>Language Testing</i> , 28(2), 221–249.	This paper reports on a survey investigating student and parents' perceptions of SBA introduced as part of the Hong Kong Certificate of Education Examinations for secondary school students. Interestingly, it found that students viewed SBA in the same way as the more formal parts of the examination and was in part determined by their parents' perceptions of SBA.	B
2011	Llosa, L. (2011). Standards-based classroom assessments of English proficiency: A review of issues, current developments, and future directions for research. <i>Language Testing</i> , 28(3), 367–382.	Based on the results of a large-scale study of school-based classroom assessment of English proficiency in the US, <b>Llosa</b> questions the usefulness of such assessments for classroom-based, formative purposes. She demonstrates that teachers are good judges of students' overall language ability for summative purposes, but their ability to judge students' mastery of individual standards is much less consistent. <b>Llosa</b> argues that teachers need a better understanding of specified standards to be able to help students achieve these standards. The challenges teachers face in operationalising standards discussed by Llosa echo those raised in earlier work by <b>DAVISON</b> (2004) and <b>CHENG ET AL.</b> (2004).	A
2012	Hill, K. & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. <i>Language Testing</i> , 29(3), 395–420.	<b>Hill and McNamara</b> adopt a bottom-up, grounded approach to developing a comprehensive framework for researching CBA. They propose a more inclusive framework that incorporates aspects of earlier models including that drawn up by <b>REA-DICKINS</b> (2001) and addresses a number of gaps in earlier research, such as the theoretical bases for observed CBA practices. Some of the conceptual points raised in this paper resonate with <b>BROADFOOT &amp; BLACK</b> (2004).	A & B
2013	Scarino, A. (2013). Language awareness literacy and self-awareness: Understanding the role of interpretation in assessment and in teacher learning. <i>Language Testing</i> , 30(3), 309–327.	Acknowledging the influence of sociocultural theories of learning on assessment practices, <b>Scarino</b> argues for an expansion of the knowledge-base of teacher-assessors. On the basis of examples from in-service training, she shows the importance of acknowledging and working with trainees' existing knowledge and preconceptions of language, and providing the tools for them to critically assess their own assessment practices to develop their assessment literacy. This paper picks up on many of the issues discussed by <b>REA-DICKINS &amp; GARDNER</b> (2000) and <b>DAVISON</b> (2004).	A
2014	Al-Nouh, N. A., Taqi, H. A., & Abdul-Kareem, M. A. (2014). EFL primary school teachers' attitudes, knowledge and skills in alternative assessment. <i>International Educational Studies</i> , 7(5), 68–84.	In a similar vein to <b>VOGT &amp; TSAGARI</b> (2014), this large-scale study investigated assessment literacy of female primary school teachers of English in Kuwait. Teachers reported that although they perceived themselves as both knowledgeable and skilful in using alternative assessment, they felt more confident and were more favourably disposed to traditional, summative testing.	B

2014	Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. <i>Language Assessment Quarterly</i> , 11(4), 374–402.	This survey explores teacher assessment literacy across seven European countries and shows that training in the field is limited and where received, tends to be limited to traditional assessments. The results, which resonate with SCARINO (2013), point to a need for more focus to be placed on language testing and assessment in pre- and in-service training courses.	B
2016	Turner, C., & Purpura, J. (2016). Learning oriented assessment in the classroom. In D. Tsigari & J. Banerjee (Eds.), <i>Handbook of second language assessment</i> (pp. 255–272). Boston, MA/Berlin, Germany: Walter de Gruyter.	In this seminal paper, <b>Turner and Purpura</b> acknowledge the complex, multifaceted nature of learner-oriented assessment (LOA), and put forward a working framework for describing the dimensions underlying LOA. Unlike much previous exploration of CBA which takes an interactional approach (e.g. REA-DICKINS, 2001; HILL & McNAMARA, 2012), their focus is more psycho-cognitive, starting with the learning process to investigate how assessment can encourage the learner and the learning process.	A
2017	Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. <i>Melbourne Papers in Language Testing and Assessment</i> , 6(1), 1–17.	<b>Hill</b> proposes a framework designed to help teachers reflect on their current assessment practices and their professional development needs. The framework is an extension of her earlier work with McNamara (HILL & McNAMARA, 2012). It takes account of the full range of assessments including more spontaneous, incidental assessment that had hitherto been largely neglected.	A
2017	Saito, H., & Inoi, Shin'ichi (2017). Junior and senior high school EFL teachers' use of formative assessment: A mixed methods study. <i>Language Assessment Quarterly</i> , 14(3), 233.	<b>Saito and Inoi</b> investigated the differential use of formative assessment among high school English-as-a-foreign language teachers in Japan. They identified three levels of formative assessment use: high, mid and low, differentiated by four strategy use formative assessment variables, namely: intentions, methods, purposes and feedback. Unlike previous studies (e.g. CHENG ET AL., 2004), Saito and Inoi found that the differences in use of formative assessment were individual and did not depend on the educational level of teachers or the type of school and did not impact on the quality of teaching. Some of the discussion in this paper touches on issues raised by HILL & McNAMARA (2012).	B
2017	Scarino, A. (2017). Developing assessment literacy of teachers of languages: A conceptual and interpretive challenge. <i>Papers in Language Testing and Assessment</i> , 6(1), 18–40.	<b>Scarino</b> argues that, with the onset of globalisation, there has been a shift away from Communicative Language Teaching to an intercultural orientation which has brought with it a need for language teachers to reconceptualise the construct to be assessed and adjust the nature of the assessment process. This has brought conceptual and interpretive challenges for teachers that have implications for teacher development of assessment literacy.	A

(Continued)

(Continued)

Year	References	Annotations	Theme (s)
2018	Bonner, S. M., Rivera, C. T., & Chen, P. (2018). Standards and assessment: Coherence from the teacher's perspective. <i>Educational Assessment and Accountability</i> , 30(1), 71–92.	Working within the US context, <b>Bonner et al.</b> investigate how teachers' personal beliefs about external tests and classroom assessment interact with their instructional alignment to externally imposed standards. Their findings suggest that the teachers' instructional use of standards, classroom assessment preferences, and beliefs about the test-driven system are moderately correlated. In line with YEUMING ET AL.'s (2006) findings, they highlight numerous inconsistencies between teacher beliefs, classroom practices and mandated policy, with teachers mimicking external assessment formats even if they are contrary to their own beliefs.	<b>B</b>
2018	East, M. (2018). Embedding the assessment into the learning: A new direction for high-stakes assessments. In S. Coffey & U. Wingate (Eds.), <i>New directions for research in foreign language education</i> (pp. 183–199). New York, NY/London, UK: Routledge.	<b>East</b> investigates the impact of the introduction of a new foreign language speaking assessment in New Zealand as part of the National Certificate of Educational Achievement. The aim of the reform is to move towards locally-based and teacher-created assessment. Although the teacher interviewees perceived the assessment as providing learning potential for students, teachers recognised a tension between the way they operationalise the assessments and the accountability necessitated by the high-stakes nature of the assessments.	<b>B</b>
2018	Leung, C., Davison, C., East, M., Evans, M., Liu, Y., Hamp-Lyons, L., & Purpura, J. (2016). Using assessment to promote learning: Clarifying construct theories and practices. In J. McE. Davis, J. M. Norris, M. E. Malone, T. H. McKay, & Y. A. Son (Eds.), <i>Useful assessment and evaluation in language education</i> (pp. 75–94). Washington, DC: Georgetown University Press.	This chapter discusses four innovative assessment practices all linked by a common theme of 'for-learning assessment' that is embedded in pedagogical practice and situated in a particular educational and sociocultural context. It highlights again the tension between teachers addressing students' needs and having to adhere to curriculum requirements, as well as the interlocking nature of the various issues that impact on the assessment such as the context, construct, nature of learning and the linguistic focus of teaching/learning.	<b>B</b>
2018	Seden, K., & Svaricek, R. (2018), Teacher subjectivity regarding assessment: Exploring English as a foreign language teachers' conceptions of assessment theories that influence student learning. <i>CEPS. Journal</i> , 8(3), 119–139.	This paper provides further evidence (see e.g. LEUNG & REA-DICKINS, 2007; YEUMING ET AL., 2006), this time from the Czech Republic, of the mismatch between assessment theory and classroom practice. <b>Seden and Svaricek</b> point out that the learning environment that teachers work in is tightly controlled by educational, institutional, social and cultural contexts all of which constrain innovative approaches to CBA, leaving teachers to continue to rely on traditional forms of assessment in the classroom.	<b>B</b>

2019	<p>Levi, T., &amp; Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. <i>Language Assessment Quarterly</i>, 17(2), 168–180.</p>	<p>Rather than relying on teacher self-report of language assessment competencies (e.g. the study by VOGT &amp; TSIGARI, 2014), <b>Levi and Inbar-Lourie</b> investigate language teachers’ assessment literacy through examining the application of their assessment knowledge following a generic course on assessment literacy. Content analysis of the assessments produced by language teachers of both English and Hebrew produced revealed their unique needs to address the multi-componential complexity of assessing language particularly in formative situations. Their work reaffirms many earlier studies highlighting the complex nature of language assessment literacy (cf. TEASDALE &amp; LEUNG, 2000; TURNER &amp; PURPURA, 2016).</p>	<b>B</b>
2020	<p>Hill, K., &amp; Ducasse, A. M. (2020). Advancing written feedback practice through teacher-researcher collaboration in a university Spanish program. In M. Poehner &amp; O. Inbar-Lourie (Eds.), <i>Toward a reconceptualization of L2 classroom assessment: Praxis and researcher-teacher partnership</i> (pp. 153–172). Cham, Switzerland: Springer Nature.</p>	<p>This chapter explores how the teacher assessment literacy resource developed by HILL (2017) can inform as well as be informed by written feedback practices employed by an experienced teacher of L2 Spanish in the Australian context. The teacher–researcher collaboration proved to be of mutual benefit, enabling the teacher to reflect in a more systematic way on her own feedback practices, and the researcher to assess and refine the resource.</p>	<b>A &amp; B</b>