

Assessment and English as a medium of instruction

Challenges and opportunities

Anna Kristina Hultgren,¹ Nathaniel Owen,² Prithvi Shrestha,¹ Maria Kuteeva,³ and Špela Mežek³

¹The Open University | ²Oxford University Press | ³Stockholm University

As English as a Medium of Instruction (EMI) continues to expand across the globe, there is a glaring absence of research on assessment. This article reviews the scarce literature to date and maps out a research agenda for the future. Drawing on Shohamy's (2001, 2007) Critical Language Testing and McNamara et al.'s (2019) notions of "fair" and "just" language assessment, our reading of the literature to date is that it has revealed considerable complexities around implementing assessment in EMI contexts, with key questions centring not only on what and who to assess but also on how and why assessment should take place. In outlining a research agenda for the future, we suggest that one way of bypassing such challenges may be to carve out a greater role for assessment for *learning* in higher education. This could capitalize on – and raise stakeholders' awareness of – bodies of knowledge that are well established within applied linguistics about the integral role of *language* in learning. Whilst we acknowledge challenges in securing institutional buy-in for putting this agenda into practice, we suggest that doing so could turn assessment challenges into opportunities and significantly enhance learning not only in EMI contexts but beyond.

Keywords: assessment, English as a Medium of Instruction, fair and just assessment, assessment for learning

Introduction

As English as a Medium of Instruction (EMI) continues to expand across the globe, there is a glaring absence of research on assessment in EMI contexts. Of course, assessment is and has always had an integral role in most educational

institutions, regardless of the language of instruction. However, as Wilkinson and Zegers imply in the first book on this topic, the new reality of EMI has not been accompanied by a consideration of the extent to which assessment ought (or ought not) to be rethought:

the fact that education takes place through a language that is not the students' mother tongue (and, in many cases, not that of the educators either) seems to have little influence on the assessment processes. Nor does the fact that students are learning both disciplinary content and language seem to influence the process. (2006, p. 30)

In the fifteen years that have passed since Wilkinson and Zegers made this observation, the lack of attention to EMI-specific assessment still largely holds true, notwithstanding important research progress in some areas.

In our contribution to this inaugural issue of the *Journal of English-Medium Instruction*, we start by reviewing the extant literature on assessment and EMI. As EMI is a global phenomenon (Macaro et al., 2018), we have sought to adopt as global an outlook as possible; however, we limit our focus to higher education, where we believe EMI has seen the fastest growth. A key theme to emerge from our review are challenges within Socio- and Applied Linguistics to constructs of key relevance: proficiency, norms, standards and even language itself, raising serious questions about how to conceptualize and implement assessment in EMI contexts. In outlining a research agenda for the future, we build on this knowledge to suggest that one way forward may be to bypass such complexities by carving out a greater role for assessment as a tool for *learning*. In particular, we see promise in approaches that highlight the role of language in learning and suggest that EMI offers opportunities for embedding this knowledge in institutional policy and practice, although we recognise significant challenges in doing so. We start, however, with a brief outline of recent theoretical developments in Socio- and Applied Linguistics.

Theoretical developments

In the past few decades, Socio- and Applied Linguistics has taken a turn to practice, increasingly emphasising fluidity, complexity and contestation of norms. As Blommaert and Backus write, globalization has brought “new social environments... characterized by *an extremely low degree of presupposability* in terms of identities, patterns of social and cultural behavior, social and cultural structure, norms and expectations” (2013, p. 13, italics in original). Such “low degree of presupposability” poses challenges for assessment which rests upon an a priori and

explicit articulation of a set of criteria against which test-takers are going to be assessed. There may be some inherent tensions, in other words, between the interpretive epistemologies underpinning recent developments in Socio- and Applied Linguistics and the predominantly positivist epistemologies upon which the field of Language Testing and Assessment is premised. In other words, if language, norms and standards are co-constructed, contextually contingent and emerge through practice in every new situation, it causes epistemological friction with assessment which is a way of measuring a time and location-bound performance of knowledge or skill, judged against some predetermined standard, and from which we can infer a claim about proficiency or make a decision about an individual test-taker.¹

EMI contexts are English as a Lingua Franca (ELF) environments in which English is not the first language of the majority of interactants (Smit, 2017). This raises questions about the appropriacy of measuring ELF users against a yardstick of an idealised English “native speaker” (the quotation marks signal the idealized status of this construct). Globalization, and the ensuing destabilization of norms, has made this “a testing time for testers” (Jenkins, 2006). As Harding and McNamara put it: “The sociolinguistic reality of English as a lingua franca (ELF) communication represents one of the most significant challenges to language testing and assessment since the advent of the communicative revolution” (2017, p. 570). Moreover, in EMI contexts “language development is not a primary intended outcome” (Pecorari & Malmström, 2018, p. 499), raising further questions around the relevance of *language* as an assessment construct and about the transferability of insights from the field of Language Testing and Assessment to EMI contexts. Although “a widely purported benefit of EMI is that it kills two birds with one stone; in other words, students simultaneously acquire both English and content knowledge” (Rose et al., 2019, p. 2), the extent to which language learning actually happens remains debated (Macaro et al., 2018).

Like most other linguistically and culturally diverse contexts, EMI settings are increasingly recognised, not as deficit environments, but more positively as environments in which a wide range of multimodal and multilingual resources can be mobilised and seen as strengths (Paulsrud, et al., 2021; Dafouz & Smit, 2020; Kuteeva et al., 2020). Moncada-Comas, writing about an EMI context, shows how students with Catalan or Spanish as their first language draw on their entire repertoire of “semiotic resources and their multimodal ensembles”

1. Such moves to social constructionism are discernible also in the wider field of Educational Assessment. Baird et al. suggest that: “[e]ducational assessment constructs are not ‘out there’ waiting to be discovered” (2017: 322). Rather, they should be seen as entities produced and interpreted by human actors, conditioned by their ideological, socio-cultural and political values. To Baird et al. (2017) assessment constructs are not (primarily) empirical but theoretical constructs.

to enable communication (2020, p. 43).² Komori-Glatz and Smit, similarly, point to a “nuanced interplay between the negotiation of meaning, repertoires ‘in flux’ and dynamic participant roles” (2021, p. 20). In other words, language proficiency cannot be meaningfully extracted from a fuller assemblage of meaning-making resources, co-constructed and negotiated by interactants. There is a clear move away, in other words, at least in some corners of the field, from “English proficiency” as autonomously conceptualized and consequently from reductionist and decontextualized understandings of “success” in EMI contexts. This inevitably broadens out the scope of assessment, posing challenges to validity. Validity is a crucial, though vexed, concept in Language Testing and Assessment, at the heart of which lies questions such as “how we can know what a test score means, and what it can be used for” (Fulcher & Davidson, 2012, p. 2). Acknowledging such inherent complexities, Bachman and Palmer, two leading scholars in the field, were early to warn against an overly reverent reliance on testing, reminding us that any test will always only reflect a part of a bigger picture (Bachman & Palmer, 1996). It is for this reason that testing is sometimes referred to as “the art of the possible” (Taylor, 2006).

On the basis of these theoretical developments, our review of the literature in the next section aligns with Critical Language Testing (Shohamy, 2001, 2017), which raises questions about the possibilities and ethics of devising valid language tests in environments where English is used as a lingua franca. We also draw on McNamara et al.’s (2019) notions of “fair” and “just” language assessment. The “fairness” of a test refers to its internal quality or *validity*. Validity, as mentioned, refers to the extent to which a test represents the construct(s) it sets out to measure (Bachman & Palmer, 2010). As research has shown, however, what exactly is an appropriate construct to measure in an EMI context is not always clear (Dimova, 2020a). The “justice” of a test, on the other hand, refers to the defensibility of the policies and values implicit in the *use* of the test: its political dimensions (McNamara et al., 2019). As McNamara et al. (2019) explain, irrespective of their validity, tests can be used in different ways for political purposes. The authors give the example of English language proficiency requirements for UK and Australian immigration policies which have fluctuated purposely in response

2. Similar moves from “deficit approaches” and an emphasis on “correctness” of form towards viewing communicative proficiency more holistically are made also in English-dominant environments. Kimura and Canagarajah, for instance, show how a South Korean Microbiologist researcher at a US university, who, despite self-assessing his English language proficiency as “limited,” makes valuable contributions to the discussion of a research group meeting by “leveraging embodied semiotic resources alongside verbal ones” (2020: 634). This leads Kimura and Canagarajah to “question institutional policies that privilege advanced proficiency in normative English grammar for skilled migrants, which leads to unfairly judging their competence” (2020: 634).

to labour shortages and surpluses. Similarly, in higher education, which is in focus here, institutions are often constrained by ministerial mandates and may be caught between conflicting incentives. While on the one hand universities may be mandated to retain and produce highly qualified graduates (which would incentivize them to set entry requirements at a high level), they may also be mandated to recruit broadly and *en masse* (which incentivizes them to set entry requirements at a lower level).

Where are we now? An overview of existing research

In this section, we review extant research into assessment in EMI contexts. Our discussion is structured into two parts. We start by discussing assessment of content lecturers (for a fuller discussion of this, see Lasagabaster, this issue). We then focus on students where at least some of the discussion has centred on entry requirements and the extent to which these would ensure students' "success" in EMI contexts. For both these groups, if there is one thing research has revealed, it is considerable complexities around devising "fair" and "just" language assessments in EMI contexts.

Assessment of content lecturers

Some research has focused on the extent to which content lecturers who do not have English as their first language are and should be assessed in order to teach through English. Partly, such questions respond to early concerns in the literature about whether the quality of teaching would be lowered by being delivered by non-native English speakers.³ Whilst it is difficult to get an overall picture of the prevalence of assessment of content lecturers, a survey by O'Dowd (2018) among 70 universities in Europe found that more than three quarters require a particular level of English proficiency from their lecturing staff. The vast majority require a threshold of what equates to B2 ("vantage or upper intermediate" in the Common European Framework of Reference (CEFR)) or CEFR C1 ("effective operational proficiency or advanced"), and a small minority require CEFR C2 ("mastery or proficiency") (O'Dowd, 2018). However, it should be noted that it is not always

3. A reviewer points out that this replicates practices in the US where in numerous states certification of the English proficiency of International Teaching Assistants (ITAs) is mandated by the state legislature in response to concerns 30 years ago or more that undergraduate students found "foreign" ITAs lacking in both intelligibility and teaching skills. The same reviewer comments: "I would hope that most of these programmes have now moved well beyond what originally motivated such laws: a narrow, native-speaker-oriented view of what intelligibility and inter-cultural communication involve."

clear how and whether such requirements are enforced in practice. Perhaps realizing that English proficiency is not the only thing of relevance for university teaching, some universities have begun to develop their own in-house procedures for assessing content lecturers, a process sometimes referred to as “certification” (Ball & Lindsay, 2013; Dimova, 2020a; Macaro et al., 2019). Whilst in-house test development is resource-intensive, it does have the advantage of tailoring the assessment to what is contextually relevant.⁴

A model example of a locally developed certification process is TOEPAS (Test of Oral English Proficiency for Academic Staff). This is a test which has been in use at the University of Copenhagen, Denmark, since 2009, but which has gradually been revised to encompass pedagogical skills in addition to English language proficiency (Dimova, 2020a; Dimova & Kling, 2018). Dimova and Kling (2018) explain how the original version of the test, which referenced formal accuracy and native-speaker norms as targets, has undergone subsequent revision in light of research that questioned the relevance of these criteria in EMI contexts. The revised test, TOEPAS 2.0, now includes descriptors related to lecturers’ abilities to present content material effectively, using summaries, examples, or emphasis of important points, as well as their ability to deliver their lectures in an organized and structured manner. Whilst some structural linguistic elements (e.g., grammar and syntax) remain among the criteria, their prominence has been reduced in the revised version of the test (Dimova, 2020a). TOEPAS also combines certification with continued professional development, offering tailored training to lecturers based on their performance in the test (Dimova, 2020c). As such, TOEPAS foregrounds assessment for *development* rather than assessment for *gatekeeping*.

The move towards contextually situated language assessment as exemplified by the revised version of TOEPAS reflects the aforementioned recognition in the field that English language proficiency is difficult to tease out from other communicative resources (see also Komori-Glatz & Smit, 2021). Drawing on McNamara et al.’s (2019) concepts of “fairness” and “justice,” assessment is “fair” when it accurately measures the construct it is meant to measure. The “justice” of the test, on the other hand, pertains to the uses of the test. Dimova (2017) reports from the University of Copenhagen that while most staff who undergo the certification process at the University of Copenhagen perceive it as a positive experience, and value the training possibilities it affords, certain staff categories may be vulnerable to the process. Temporary staff, for instance, may fear that failing the test limits their prospects of permanent employment. International staff, in turn, who do

4. This replicates practices in some major US universities who also conduct their own tailored assessments of the proficiency of International Teaching Assistants (ITAs), rather than relying on a standardized oral test like the SPEAK. The assessment is typically linked to teaching support for prospective ITAs in the form of credit courses or workshops.

not have Danish as their first language, may fear that failing the test would preclude them from teaching altogether as they cannot, like their Danish-speaking colleagues, fall back on teaching in Danish (Dimova, 2017). This demonstrates why Klaassen and Räsänen refer to assessment of the language proficiency of university teaching staff as “a delicate issue” (2006, p. 236), and why Dimova (2017) recommends designing the test in consultation with unions and higher education policy makers.

Student entry requirements

Research exploring student entry requirements has documented considerable variation in the extent to which universities require proof of English language proficiency in order to be admitted to EMI programmes (Galloway, 2020; Stuers, 2019; Wächter & Maiworm, 2014).⁵ Some universities require proof from established test providers (TOEFL, IELTS, etc.) whereas others appear to infer English language proficiency on the basis of a school-leaving certificate or previous English-medium schooling (Galloway, 2020; Dimova, 2020b; Stuers, 2019; Wächter & Maiworm, 2014; Wilkinson & Zegers, 2006). Some appear to go to extra lengths to ensure students’ preparedness to study in English, making use of entrance exams and preparatory courses which may be combined with general course preparation (Galloway, 2020; Wächter & Maiworm, 2014). It is not always clear if proof of English language proficiency applies to international or domestic students or to both, but in general native users of English tend to be exempt (Stuers, 2019; Dimova, 2020b; Wächter & Maiworm, 2014). Sometimes, national policies determine what is possible and not. In the Netherlands, for instance, laws preclude the use of English language proficiency requirements for domestic students on the grounds that no requirements in addition to those already in place can be used; however, they are allowed for international students (Wilkinson & Zegers, 2006). In terms of cut-scores, the average across Europe is an equivalent of CEFR B2, and a range of both international and local tests are accepted as proof hereof (Wächter & Maiworm, 2014; Stuers, 2019; Klaassen & Räsänen, 2006). Worldwide, cut-scores range from what equates to CEFR B1 to C1 levels

5. Such variation is not particularly surprising as universities and nations are known to vary widely in their general matriculation processes. To gain entry to some graduate programmes in the US, e.g., international candidates may need no less than three different proofs for admission: international English language test, subject knowledge test for the chosen area of study and a GRE (Graduate Record Examinations) score showing aptitude for verbal reasoning, quantitative reasoning, analytical writing, and critical thinking skills.

(Galloway, 2020). In sum, no uniform picture emerges regarding student entry requirements.

Some studies have sought to establish the extent to which English language entry requirements provide a valid predictor for academic success. Such studies have tended to yield conflicting results which may reflect inherent difficulties in equating EMI “success” with a numerical score (Owen et al., 2021; Curle et al., 2020; Rose et al., 2019). Where some research seeks answers to which cut-score are optimal to ensure academic success, others argue that such questions do not reflect the complexity and situated nature of communication and learning (Komori-Glatz & Smit, 2021). We detect a slight tension in the literature between, on the one hand, positivist orientations that seek evidence-based answers to what constitutes “success” in EMI and, on the other, interpretivist orientations that believe such questions are not helpful and can never be conclusively answered. Doubts over the relevance of English-language entry requirements as determinants of success are echoed by those closest to students. A lecturer at a university in Denmark report: “I’m not sure those tests are enough to say that a student is good at writing scientific English, or good enough at communicating” (Dimova, 2020b, p.10). Many lecturers appear to see English language entry requirements mainly as a bureaucratic exercise with little or no relevance once the student has gained entry (Dimova, 2020b; see also Breeze, 2012).

To sum up, research has revealed considerable complexities with regard to which constructs are relevant to assess in EMI contexts, both with regard to the assessment of lecturers and students. In McNamara et al.’s (2019) terminology, there are questions over how to develop “fair” assessment in EMI contexts that measures what is relevant. With regard to lecturer certification, the revised version of TOEPAS, developed in-house at the University of Copenhagen, is often hailed as a model in that it co-thinks language proficiency with pedagogical skills while seeking to move away from “English native speaker proficiency” as a norm (Dimova, 2020a). With regard to student entry requirements, research has revealed a picture in flux with key questions centring on whether English language entry requirements would offer a “fair” assessment of students in the sense of accurately predicting their academic success. There are also questions around what, if anything, English language proficiency requirements would bring that is not already catered for by any gatekeeping requirements already in place, such as general admission requirements for students and teaching qualifications for lecturers.⁶

6. A survey of Europe shows that countries vary in their requirements for faculty to hold general teaching qualifications (Gaebel & Zhang, 2018). Similarly, student matriculation requirements are known to vary widely between countries and between universities.

Given such uncertainties around what constitutes “fair” assessment in EMI contexts, further questions arise as to how to undertake “just” assessment. Cut-scores for entry requirements tend to be set arbitrarily, notwithstanding an often “blind faith” by admission officers in their reliability and appropriateness (Stuers, 2019). The way in which admission tests are administered in practice also tends to be “determined by pragmatism and compromise” (Deygers & Malone, 2019, p.347). As mentioned, universities are often constrained by ministries which is why entry requirements can fluctuate in accordance with domestic policies on immigration and higher education policies on student uptake and output. Such issues all relate to the uses of tests (what McNamara et al., 2019, refer to as the “justice” of a test) and demonstrate that assessment is more than scientific measures of validity and in need of being conceptualized within a wider socio-political context.

Moreover, if “English language proficiency,” as autonomously defined, turns out to be a relatively poor predictor of teaching competence for lecturers and academic success for students, then it is not obvious that assessment for gatekeeping purposes should differ for native and non-native users of English, but in many cases they continue to do so. This is in spite of long-standing calls from Language Testing and Assessment scholars and others for this dichotomy to be dismantled (see also Elder & Davies, 2006; Harding & McNamara, 2017; Jenkins & Leung, 2019; Kuteeva, 2020; McNamara et al., 2019; Shohamy, 2001, 2017; Taylor, 2006;):

the growing awareness of the nature of English as a lingua franca communication overturns all the givens of the communicative movement as it has developed over the last 30 or 40 years. The distinction between native and non-native speaker competence, which lies at the heart of the movement, can no longer be sustained; we need a radical reconceptualization of the construct of successful communication that does not depend on this distinction. (McNamara, 2014, p. 21)

Pecorari and Malmström (2018) have similarly argued that there is no straightforward way of differentiating between EMI and non-EMI contexts. There may well turn out to be as much variation in language development needs *within* each context as *between* them.

Where should we go next: Assessment for learning

In moving forward, we suggest that research could usefully carve out a greater space for how assessment can be used to support and enhance learning. A lesser focus on English language proficiency as autonomously conceived in favour of a greater focus on communication through a range of semiotic, multimodal and multilingual resources, as advocated in much of the literature, therefore seems an

appropriate way forward. Crucially, this approach would be less concerned with whether a student's or lecturer's proficiency in English is of a certain "standard" (a vexed concept) and more with how a greater awareness of language can support learning and teaching. The emphasis is on communication rather than on a specific language in line with the idea that "named languages do not neatly correspond to the mental representations nor the language practices of racialized bilingual speakers (or of any speakers)" (García et al., 2021, p.16).⁷

The point about emphasising assessment for *learning* echoes one made by Baird et al. (2017) within the broader field of Educational Assessment. These authors take aim at international tests such as PISA, which compare and rank national education systems against one another. The authors argue that the international testing regimes have made the field of Educational Assessment overly concerned with mechanistic matters such as test constructs and validity to the neglect of how such constructs serve to promote and enhance learning. Their article, titled "Assessment and learning: Fields apart?" warns about the dangers of assessment becoming a tool for benchmarking to the detriment of it serving as a tool for learning. Similar criticism has long been voiced within the field of Language Testing and Assessment, particularly around the "fairness" and "justice" of using standardised English language proficiency tests such as IELTS and TOEFL for gatekeeping purposes in higher education, immigration and professional registration internationally (Elder & Davies, 2006; Harding & McNamara, 2017; Jenkins & Leung, 2019; McNamara et al., 2019; Shohamy, 2001, 2017; Taylor, 2006).

A wealth of knowledge from within Applied Linguistics already exists on how to enhance learning in EMI contexts by raising awareness of the role of language. Some such bodies of knowledge include Content and Language Integrated Learning, Language Awareness, Academic Literacies, Systemic Functional Linguistics, Disciplinary Literacy, New Literacy Studies, English for Academic Purposes, Dynamic Assessment, and many others. We are less concerned with the differences between these fields than their commonalities, namely their aim to render visible the role of language in promoting contextually appropriate learning. However, whilst the benefits of such bodies of knowledge are well-known to linguists, they are less so to key stakeholders in EMI contexts, including policy makers, content lecturers and students.

To give a more concrete idea of what we are suggesting, we conclude this section by giving a few examples of how a language-sensitive approach to assessment for learning has been implemented in practice in different places across the

7. By "racialized bilinguals," García et al. mean "people who, as a result of long processes of domination and colonization, have been positioned as inferior in racial and linguistic terms" (2021: 2), a label that may describe some EMI contexts more aptly than others.

world, both English-dominant and non-English-dominant (see also Lasagabaster & Doiz, 2021). Given our assertion that there is nothing to inherently distinguish language development needs in EMI from those in non-EMI contexts, there is arguably a lot of knowledge exchange that could usefully be engaged in between these settings. In English-dominant contexts, many of these initiatives have been going on for well over 20 years or more whereas in EMI contexts they may be more recent. On the other hand, EMI contexts often render visible language development needs that may be more hidden in English-dominant contexts.

South Africa

Jacobs (2006) shows how academic literacy can be incorporated into assessment practices by having content and language teachers work together in assessing students' written work in a Civil Engineering course. The collaboration consisted in the content lecturer and the language specialist devising the assessment task collaboratively. They then independently assessed it, focusing, in turn, on each of their areas of expertise: language and composition in the case of the language specialist and content in the case of the content lecturer. By marking the same copy in turn, each assessor was exposed to the feedback of the other party and, in addition to marking the script, the exercise thereby served the additional purpose of professional development by sensitising both teachers to how language and content could be integrated into the more general assessment practices in the department.

United Kingdom

Although not an EMI context by most definitions, much work has been done in the UK on academic literacies and on how attention to language in assessment can promote learning. Shrestha's work (Shrestha, 2020; Shrestha & Parry, 2019) focused on designing academic literacy activities and assessment with early childhood lecturers at The Open University, each contributing their share. Shrestha, as an academic literacy specialist, worked collaboratively with a team of early childhood academics to design first and second year core modules for early childhood studies in which academic literacy activities were embedded. Each assignment was linked with relevant academic literacy activities, thus making language and academic literacy visible to students and making them aware of their academic communication. The assignments were marked by Associate Lecturers, who, although not involved in devising the assessment, were prompted by the marking rubric to pay attention not only the subject knowledge but also to academic literacy. Thus, as in Jacob's example, by having academic literacy specialists, early childhood lecturers and Associate Lecturers work collaboratively together, each

was sensitized to the intertwined nature of language and content, and ultimately, this benefited the students as well by enhancing both their academic literacy as well as their understanding of the subject knowledge.

Sweden

Kuteeva, a co-author of this article, provides the example of the first full EMI programmes at Stockholm University School of Business which wanted to make language explicit and had English integrated in all four content modules offered in the first semester. The preparation of these four content modules involved collaboration between the EMI programme director, content lecturers, and the Centre for Academic English. The learning outcomes for each content module included an item concerning the use of English in the context of that discipline, with a pass/fail assignment in Academic English as a requirement. The contents of Academic English seminars and the final assignment were developed in collaboration with content lecturers and graded by language instructors. This worked reasonably well, although academic English lecturers reported that some students thought they already knew more than enough English. Most students, however, would see the point, at least for the initial modules. This practice was in place for a few years between 2012 and 2017 but was then scrapped during a restructuring process. Ironically, this was done through a top-down decision by the department responsible for providing these language courses, as they were considered too labour-intensive and not central to its core activities. Later, even some of the EMI programmes were also redesigned and rebranded, and they no longer include any explicit language instruction. This case underscores a need for specific resource allocation and institutional commitment to integrating language and content in EMI programmes.

These are only three among a wealth of grassroot examples across the world that integrate language and content in assessment in ways that might support and enhance learning.⁸

Despite many promising grassroot examples of how a greater awareness of language can support and enhance learning, experience has shown that such knowledge often ends up being under-utilised because of a lack of institutional buy-in and resource commitment.

8. We are grateful to a reviewer for pointing out that some English-dominant contexts make use of post-entry language assessment (PELA) as a way to diagnose and develop students' academic language proficiency post-admission (see, e.g., Humphreys, 2017; Read, 2015; Murray, 2015).

How do we put it into practice: Securing institutional buy-in

Although there is wide agreement, at least among linguists, about the value of promoting and raising awareness of the role of language in learning, there are often considerable challenges implementing such initiatives in practice. Certainly, there are many successful grassroots activities (see, e.g., Lasagabaste & Doiz, 2021; Block & Mancho-Barés, 2020; Wilkinson & Zegers, 2006); however, it emerges from the literature that to keep them going on a sustained basis for more than a few years or on a large scale (involving more than one course) is very difficult. As Kuteeva's example above illustrates, research, policy and practice on assessment for learning in EMI contexts is unlikely to make substantial headway in practice unless it secures institutional buy-in. The resistance from subject lecturers to becoming "language teachers" is a major factor, as is the lack of institutional buy-in and complex organizational structures which make it difficult to collaborate across disciplinary and departmental boundaries.

At tertiary level, educators are by nature highly specialised in their field, whether this is content lecturers, language specialists, educational developers, and so on. No one individual is likely to be in a position to devise content and language-integrated assessments. This is different from pre-tertiary level where content and language integrated learning is perhaps more readily incorporable into assessment (deBoer & Leontjev, 2020). In higher education contexts, however, EMI content lecturers typically do not see themselves as language teachers (Airey, 2012; Block & Mancho-Barés, 2020), and language specialists are themselves highly specialised in their field. The question therefore becomes how to engage in the requisite multidisciplinary collaboration given that universities are complex, rigid and resource-constrained organizational structures. Experience has shown that it takes a heroic commitment from a dedicated and talented team of subject and language specialists to make it work, even on a short-term basis.⁹ Research on language-sensitive assessment for learning is therefore unlikely to make substantial headway in practice unless it sufficiently addresses the following question:

- How do we secure the institutional buy-in from key stakeholders to implement language-sensitive assessment for learning in higher education?

Faced with this challenge, it is worth reminding ourselves that embedding content and language integrated learning into assessment need not be about reinventing the wheel. Despite often overt denials by content lecturers that they are language teachers, it is not uncommon for language-related components to feature as part

9. We owe this point to a reviewer and it corresponds to our own experience.

of existing assessment practices even in fields that are not directly concerned with language. Students of zoology and physics, e.g., may well be assessed on how their assignment is written, presented or structured. Block and Mancho-Barés (2020), e.g. show how lecturers in a veterinary course at a Spanish university, despite denying that they are language teachers, do not shy away from assessing their students on their oral presentational skills in English in addition to their content understanding. A realistic and practical way forward, then, may be to capitalize on the assessment structures already in place, but granting language a more prominent place within them, by facilitating, as needed, greater collaboration between language experts and subject experts. Crucially, it will also be necessary to persuade those who hold the purse strings that such collaboration can support and promote institutional and educational goals to enhance learning and attainment.

We suggest that assessment is a good place to start when it comes to raising policy makers' awareness of the role of language in learning and educational attainment. Assessment can help to set and clarify educational priorities and it can play a key motivational role for learners, teachers and institutions. Assessment determines what knowledge and skills should be prioritised and valued in educational contexts, and beyond that, in society itself. In the case of summative assessment, it defines what aspects of learning will formally be given credit. Assessment, in short, has the potential to significantly influence practice and, in turn, to impact learning (Baird et al., 2017). Thus, we would propose that assessment is a promising locus for raising stakeholders' awareness of the value of language-sensitive pedagogy and its potential for enhancing learning in EMI contexts, and beyond.

Conclusion

This article has reviewed literature on assessment in EMI contexts and identified considerable complexities around devising "fair" and "just" assessment. As a way forward, we have suggested reconceptualizing assessment as a tool for enhancing learning by emphasising and raising stakeholders' awareness of the important role of language. We have discussed the challenges that need to be overcome in order to put this agenda into practice. Whilst often constrained by a lack of resources, institutional decision makers may be persuaded by evidence that greater attention to language in assessment can support and promote educational outcomes and institutional priorities. Beyond that, setting assessment priorities and firmly linking them up with teaching and learning could promote reflections around the purpose and role of universities in contemporary society and about the kinds of graduates we want to emerge from our higher education institutions; graduates who are fit not only for employment but also informed, engaged, reflective and

socially responsible individuals fit to navigate an increasingly complex globalized world.

Funding

The corresponding author acknowledges UK Research and Innovation for funding under a Future Leaders Fellowship (Grant reference: MR/T021500/1). All authors acknowledge Educational Testing Service for funding the research partially underpinning this work under a TOEFL® research grant.

Acknowledgements

We thank two anonymous and two named reviewers (Slobodanka Dimova and Joyce Kling) for their insightful comments. Finally, our thanks go to the journal editors for initiating the journal and this special issue.

References

- Airey, J. (2012). I don't teach language: The linguistic attitudes of physics lecturers in Sweden. *ALLA Review*, 25, 64–79. <https://doi.org/10.1075/aila.25.05air>
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford University Press.
- Baird, J.-A., Andrich, D., Hopfenbeck, T.N., & Stobart, G. (2017). Assessment and learning: fields apart? *Assessment in Education: Principles, Policy & Practice*, 24(3), 317–350. <https://doi.org/10.1080/0969594X.2017.1319337>
- Ball, P., & Lindsay, D. (2013). Language demands and support for English-medium instruction in tertiary education. Learning from a specific context. In A. Doiz, D. Lasagabaster, & J.M. Sierra (Eds.), *English-medium instruction at universities* (pp. 44–62). Multilingual Matters.
- Block, D., & Mancho-Barés, G. (2020). NOT English teachers, except when they are: The curious case of oral presentation evaluation rubrics in an EMI-in-HE context. In D. Block, & S. Khan (Eds.), *The secret life of English-medium instruction in higher education: Examining microphenomena in context* (pp. 96–119). Routledge. <https://doi.org/10.4324/9781003005667-5>
- Blommaert, J., & Backus, A. (2013). Superdiverse repertoires and the individual. In I. Saint-Georges & J.-J. Weber (Eds.), *Multilingualism and multimodality: The future of education research*. Sense Publishers. https://doi.org/10.1007/978-94-6209-266-2_2
- Bowles, H., & Murphy, A. C. (Eds.). (2020). *English-medium instruction and the internationalization of universities*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-47860-5>
- Breeze, R. (2012). *Rethinking academic writing pedagogy for the European university*. Brill. <https://doi.org/10.1163/9789401207959>
- Curle, S., Yuksel, D., Soruç, A., & Altay, M. (2020). Predictors of English medium instruction academic success: English proficiency versus first language medium. *System*, 95. <https://doi.org/10.1016/j.system.2020.102378>

- Dafouz, E., & Smit, U. (2020). *ROAD-MAPPING English medium education in the internationalised university*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-23463-8>
- deBoer, M., & Leontjev, D. (Eds.) (2020). *Assessment and learning in content and language integrated learning (CLIL). Classrooms approaches and conceptualisations*. Springer. <https://doi.org/10.1007/978-3-030-54128-6>
- Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, 36(3), 347–368. <https://doi.org/10.1177/0265532219826390>
- Dimova, S. (2017). Life after oral English certification: The consequences of the test of oral English proficiency for academic staff for EMI lecturers. *English for Specific Purposes*, 46, 45–58. <https://doi.org/10.1016/j.esp.2016.12.004>
- Dimova, S. (2020a). Language assessment of EMI content teachers: What norms. In M. Kuteeva, K. Kaufhold, & N. Hynninen (Eds.), *Language perceptions and practices in multilingual universities* (pp. 351–378). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-38755-6_14
- Dimova, S. (2020b). English language requirements for enrolment in EMI programs in higher education: A European case. *Journal of English for Academic Purposes*, 47(1–13), <https://doi.org/10.1016/j.jeap.2020.100896>
- Dimova, S. (2020c). The role of feedback in the design of a testing model for social justice. *Journal of Contemporary Philology*, 3(1), 21–34. <https://doi.org/10.37834/JCP2030021d>
- Dimova, S., & Kling, J. (2018). Assessing EMI lecturer language proficiency across disciplines. *TESOL Quarterly*, 52(3), 534–656. <https://doi.org/10.1002/tesq.454>
- Elder, C., & Davies, A. (2006). Assessing English as a Lingua Franca. *Annual Review of Applied Linguistics*, 26, 282–304. <https://doi.org/10.1017/S0267190506000146>
- Fulcher, G. & Davidson, F. (2012). Introduction. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 1–18). Routledge.
- Gaebel, Michael, & Zhang, T. (2018). *Trends 2018. Learning and teaching in the European higher education area*. European University Association.
- Galloway, N. (2020). *English in higher education – English medium Part 1: Literature review*. British Council.
- García, O., Flores, N., Seltzer, K., Li, W., Otheguy, R., & Rosa, J. (2021). Rejecting abyssal thinking in the language and education of racialized bilinguals: A manifesto. *Critical Inquiry in Language Studies*. <https://doi.org/10.1080/15427587.2021.1935957>
- Harding, L., & McNamara, T. (2017). Language assessment. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca*. Routledge. <https://doi.org/10.4324/9781315717173-46>
- Humphreys, P. (2017). EMI in Anglophone nations: Contradiction in terms or cause for consideration? In B. Fenton-Smith, P. Humphreys, & I. Walkinshaw (Eds.), *English medium instruction in higher education in Asia-Pacific* (pp.93–114). Springer. https://doi.org/10.1007/978-3-319-51976-0_6
- Jacobs, C. (2006). Integrated assessment practices: When language and content lecturers collaborate. In R. Wilkinson, V. Zegers, & C. van Leeuwen (Eds.), *Bridging the assessment gap in English-medium higher education* (pp. 141–158). AKS-Verlag.
- Jenkins, J. (2006). The spread of EIL: A testing time for testers. *ELT Journal*, 60(1), 42–50. <https://doi.org/10.1093/elt/ccio80>
- Jenkins, J., & Leung, C. (2019). From mythical 'standard' to standard reality: The need for alternatives to standardized English language tests. *Language Teaching*, 52(1), 86–110. <https://doi.org/10.1017/S0261444818000307>

- Kimura, D., & Canagarajah, S. (2020). Embodied semiotic resources in research group meetings: How language competence is framed. *Journal of Sociolinguistics*, 24(5), 634–655. <https://doi.org/10.1111/josl.12435>
- Klaassen, R. G., & Räsänen, A. (2006). Assessment and staff development in higher education for English-medium instruction: A question-raising article. In R. Wilkinson, V. Zegers, & C. van Leeuwen (Eds.), *Bridging the assessment gap in English-medium higher education* (pp. 235–255). AKS-Verlag.
- Komori-Glatz, M., & Smit, U. (2021). Exploratory interactive explaining (EXINTEX): Constructing disciplinary knowledge in two multilingual university settings. *Applied Linguistics*. <https://doi.org/10.1093/applin/amab023>
- Kuteeva, M. (2020). Revisiting the ‘E’ in EMI: Students’ perceptions of standard English, lingua franca, and translanguaging practices. *International Journal of Bilingualism and Bilingual Education*, 23(3), 287–300. <https://doi.org/10.1080/13670050.2019.1637395>
- Kuteeva, K. Kaufhold & N. Hynninen (Eds.). (2020). *Language perceptions and practices in multilingual universities*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-38755-6>
- Lasagabaster, D. (2022). Teacher preparedness for English-medium instruction. *Journal of English-Medium Instruction* 1(1), 48–65.
- Lasagabaster, D., & Doiz, A. (Eds.). (2021). *Language use in English-medium instruction at university: International perspectives on teacher practice*. Routledge. <https://doi.org/10.4324/9781003134534>
- Macaro, E., Curle, S., Pun, J., An, J. & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36–76. <https://doi.org/10.1017/S0261444817000350>
- Macaro, E., Jiméñez Muñoz, A., & Lasagabaster, D. (2019). The importance of certification of English medium instruction teachers in higher education in Spain. *Porta Linguarum*, 32(June), 103–118.
- McNamara, T. (2014). 30 years on – Evolution or revolution? *Language Assessment Quarterly*, 11, 226–232. <https://doi.org/10.1080/15434303.2014.895830>
- McNamara, T., Knoch, U., & Fan, J. (2019). *Fairness, justice, and language assessment: The role of measurement*. Oxford University Press.
- Moncada-Comas, B. (2020). “Being a student” and “doing education.” In D. Block & S. Khan (Eds.), *The secret life of English-medium instruction in higher education: Examining microphenomena in context* (pp. 43–69). Routledge. <https://doi.org/10.4324/9781003005667-3>
- Murray, N. (2015). *Standards of English in higher education*. Cambridge University Press.
- O’Dowd, R. (2018). The training and accreditation of teachers for English medium instruction: An overview of practice in European universities. *International Journal of Bilingual Education and Bilingualism*, 21(5), 553–563. <https://doi.org/10.1080/13670050.2018.1491945>
- Owen, N., Shrestha, P. N., & Hultgren, A. K. (2021). Researching academic reading in two contrasting English as a medium of instruction contexts at a university level. *ETS Research Report Series Research Report*. <https://doi.org/10.1002/ets2.12317>
- Paulsrud, BA., Tian, Z., & Toth, J. (2021). *English-medium instruction and translanguaging. Multilingual Matters*.
- Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and English medium instruction. *TESOL Quarterly*, 52(3), 496–515. <https://doi.org/10.1002/tesq.470>

- Read, J. (2015). *Assessing English proficiency for university study*. Palgrave Macmillan. <https://doi.org/10.1057/9781137315694>
- Rose, H., Curle, S., Aizawa, I., & Thompson, G. (2019). What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation. *Studies in Higher Education*, 1–13. <https://doi.org/10.1080/03075079.2019.1590690>
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson.
- Shohamy, E. (2017). ELF and critical language testing. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca*. Routledge. <https://doi.org/10.4324/9781315717173-47>
- Shrestha, P. (2020). *Dynamic assessment of students' academic writing: Vygotskian and systemic functional linguistic perspectives*. Springer. <https://doi.org/10.1007/978-3-030-55845-1>
- Shrestha, P. & Parry, J. (2019). The effect of embedded academic literacy activities on student writing in a distance learning module. In *BALEAP Conference: Innovation, exploration and transformation*, 12–14 April 2019, University of Leeds, BALEAP.
- Smit, U. (2017). Beyond monolingualism in higher education. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca*. Routledge. <https://doi.org/10.4324/9781315717173-32>
- Stuers, S. (2019). Access to and experience of English-medium instruction in higher education in Germany: A study into English language entry requirements and their relevance (Unpublished EdD dissertation). The Open University.
- Taylor, L. (2006). The changing landscape of English: implications for language assessment. *ELT Journal*, 60(1), 51–60. <https://doi.org/10.1093/elt/ccio81>
- Wächter, B., & Maiworm, F. (2014). *English-taught programmes in European higher education ACA papers on international cooperation in Education: The state of play in 2014*. ACA Papers on International Cooperation in Education.
- Wilkinson, R., & Zegers, V. (2006). Assessing incipient linguistic competences: An institutional perspective. In R. Wilkinson, V. Zegers, & C. van Leeuwen (Eds.), *Bridging the assessment gap in English-medium higher education* (pp. 61–76). AKS-Verlag.

मूल्यांकन र अंग्रेजी माध्यमको शिक्षा: चुनौतीहरू र अवसरहरू

सार (Nepali abstract)

अङ्ग्रेजी माध्यमको शिक्षा (EMI) को विश्वभर विस्तार हुँदै गए पनि, EMI मा मूल्याङ्कन अनुसन्धानको स्पष्ट अभाव देखिन्छ। यस लेखले हाल सम्मको दुर्लभ यो क्षेत्रमा भएको अनुसन्धानको समीक्षा गर्दछ र भविष्यको लागि अनुसन्धान एजेन्डा नक्सा कोर्दछ। शोमामीको (२००९, २००७) 'क्रिटिकल लैंग्वेज टेस्टिंग' र म्याकनामारा र अन्य (२०१९) को "निष्पक्ष" र "न्याय पूर्ण" भाषा मूल्याङ्कनका धारणालाई मध्य नजर राख्दा, आजसम्मको अनुसन्धानको समीक्षाले EMI मा मूल्याङ्कन कार्यान्वयन गर्ने क्रममा धेरै जटिलताहरू प्रकट गरेको छ | साहित्यको हाम्रो पढाइले मूल्याङ्कन कार्यान्वयनमा धेरै जटिलताहरू प्रकट गरेको छ। EMI सन्दर्भहरूमा, मुख्य जटिल प्रश्नहरू के र कसलाई मूल्याङ्कन गर्नुपर्छ भन्नेमा समेत केन्द्रित हुनुपर्ने देखिन्छ। भविष्यको लागि अनुसन्धान एजेन्डाको रूपरेखामा, कोनको लागि, हामी सुझाव दिन्छौं कि त्यसता चुनौतिहरूलाई सामना गर्ने एउटा तरिका उच्च शिक्षामा 'सिकाइको लागि मूल्याङ्कन' लाई ठूलो भूमिका निर्वाह गर्नु दिनु हो। यसले भाषाविज्ञानमा भाषाको राम्ररी स्थापित भएको ज्ञान या मान्यता- सिकाइमा भाषाको अभिन्न भूमिकाको- र सरोकारवालाहरूको चेतना बढाउन सक्छ। यस एजेन्डालाई व्यवहारमा उतार्नका लागि संस्थागत सहमती

प्राप्त गर्ने चुनौतीहरू स्वीकार गर्दा, हामी सुझाव दिन्छौं कि त्यसो गर्दा मूल्याङ्कन चुनौतीहरू परिणत गर्न सकिन्छ र EMI भित् मात्र नभई बाहिर सिकाइलाई उल्लेखनीय रूपमा वृद्धि गर्न सकिन्छ।

Address for correspondence

Anna Kristina Hultgren
The Open University
Stuart Hall Building
Milton Keynes MK6 7AA
UK
kristina.hultgren@open.ac.uk

Co-author information

Nathaniel Owen
Oxford University Press
nathaniel.owen@oup.com
Prithvi Shrestha
The Open University
prithvi.shrestha@open.ac.uk

Maria Kuteeva
Department of English
Stockholm University
maria.kuteeva@english.su.se
Špela Mežek
Department of English
Stockholm University
spela.mezek@english.su.se

Publication history

Date received: 1 September 2021
Date accepted: 28 October 2021