

Understanding multilingualism in internationalised higher education institutions: perspectives on fair assessment

Ingrid de Saint-Georges

To cite this article: Ingrid de Saint-Georges (03 Dec 2025): Understanding multilingualism in internationalised higher education institutions: perspectives on fair assessment, International Journal of Multilingualism, DOI: [10.1080/14790718.2025.2592120](https://doi.org/10.1080/14790718.2025.2592120)

To link to this article: <https://doi.org/10.1080/14790718.2025.2592120>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 03 Dec 2025.



Submit your article to this journal [↗](#)



Article views: 182



View related articles [↗](#)



View Crossmark data [↗](#)

Understanding multilingualism in internationalised higher education institutions: perspectives on fair assessment

Ingrid de Saint-Georges

Department of Humanities, University of Luxembourg, Esch-sur-Alzette, Luxembourg

ABSTRACT

This study investigates how multilingualism and internationalisation intersect to shape students' experiences of assessment in higher education. Focusing on a trilingual Master's program, the research draws on semi-structured interviews from a co-research project with students. The analysis identifies three dominant assessment scripts that challenge fairness: assumptions of monolingual proficiency, stable competence across communicative modes, and shared familiarity with academic task types. These scripts often disalign with the diverse linguistic repertoires and educational trajectories of plurilingual students, leading to perceived inequities in assessment. However, the study also highlights pedagogical adjustments – such as integrating students' profiles, scaffolding the assessment journey, and centering content over form – that students perceive as supportive. Framed through the lens of 'assessment as cultural script', the findings call for more reflexive, inclusive assessment practices that acknowledge the linguistic and sociocultural realities of internationalised classrooms. The paper contributes to ongoing debates on equity in higher education and advocates for culturally and linguistically responsive assessment design.

ARTICLE HISTORY

Received 18 October 2024
Accepted 13 November 2025

KEYWORDS

Multilingualism; evaluation;
higher education;
internationalisation;
assessment; fairness

Introduction

In recent years, the linguistic and cultural landscapes of higher education have undergone significant transformation. As students' mobility increases¹ and classrooms become more international (Knight, 2017), questions around fairness in assessment have gained renewed urgency. Yet, despite these shifts, many assessment practices have continued operating on assumptions that students' cohorts constitute linguistically and culturally homogeneous groups (Shohamy, 2011), overlooking how multilingualism² interact with internationalisation to affect students' learning (Hu, 2018) and the consequences of internationalisation for fair assessment.

This paper explores how such assumptions of linguistic homogeneity and shared academic norms play out in practice by examining the experiences of students enrolled

CONTACT Ingrid de Saint-Georges  Ingrid.desaintgeorges@uni.lu  Department of Humanities, University of Luxembourg, Maison des Sciences Humaines, 11 Porte des Sciences, Esch-sur-Alzette L-4366, Luxembourg

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

in a trilingual Master's program at the University of Luxembourg. Based on qualitative data from a co-research project conducted with students (Clark et al., 2022), the study investigates how learners perceive assessment in a multilingual, internationalised classroom, by asking: what implicit assessment scripts shape students' experience? Do these scripts create tensions or barriers to demonstrating their knowledge? And what kind of pedagogical adjustments do students perceive as supportive?

Drawing on the notion of 'assessment as cultural script' (Elwood & Murphy, 2015), the paper argues that fairness in evaluation must be understood as a socially and culturally situated process – one that intersect with students' linguistic repertoires, educational trajectories, and understanding of academic norms. By investigating the implicit scripts that shape assessment, the research explores how these scripts align – or clash – with the lived realities of plurilingual learners in one programme. By foregrounding students' voices, the study contributes to ongoing conversations about equity in higher education and responds to growing calls for more reflexive, inclusive approaches to assessment design in multilingual learning environments across levels of education (De Backer et al., 2017; Menken & Shohamy, 2015).

Fairness in assessment

The evolution of fairness in educational assessment

Evaluation and assessment are consequential in Higher Education: the results learners get on assessment tasks often carry important weight on a learner's educational path. Evaluations, moreover, come in various forms (tests, oral exams, essays, portfolios, self-assessment and more) and serve multiple purposes – including learning, selection, certification, and communication about student performance (Anderson & Rogan, 2010, p. 51). Given these stakes, it is not surprising that researchers and educational stakeholders have continuously invested efforts in monitoring, and improving the quality of evaluation.

In these efforts to improve the quality of evaluation, one dimension has played a central role: the quest for fairness. In today's context, fairness is often associated with the principle of equity – to make assessment fair, instructors must ensure that all students have just and meaningful opportunities to demonstrate their learning, regardless of their backgrounds (Gipps & Stobart, 2009; Lam, 1995; Stobart, 2005). Equity, however, does not imply identical outcomes. Rather, 'it asserts that assessment practice and interpretation of results need to be fair and just for all groups' (Gipps & Stobart, 2009, p. 106) and that assessment practices be responsive to the diverse needs of learners.

Over the past decades, many innovations in assessment have emerged from this evolving understanding of fairness – not as uniformity, but as responsiveness to difference (Secada, 1989; Spolsky, 2008; Stobart, 2005; Tierney, 2016). Initially, for example, standardised testing emerged in part as a critique of previously existing selections based on right-of-birth or other merit schemes (Gipps & Stobart, 2009, p. 107) that seemed to favour certain groups over others. It became a pivotal part of most educational cultures as an appeal to fairness.

This technical view of fairness – focused on psychometric validity and reliability (AERA Standards, 2014), however, was later critiqued for overlooking the social and contextual dimensions of assessment.

In the 1960s, inspired by the civil rights movements in the US, a wave of researchers began to consider that standardised educational failed to consider the relational and contextual nature of learning and assessments. This tradition started to question whether assessments designed for hegemonic learners might disadvantage other groups (e.g. minority ethnic groups; learners from different social classes) (Tierney, 2016). Researchers began to emphasise the importance of interpreting fairness as a socially situated concept – one that varies across cultural and demographic groups. This perspective called for attention not only to the assessment itself, but also the broader conditions that shape it, such as students' access to resources and opportunities beforehand, as well as the social implications of how assessment outcomes are interpreted and used (Stobart, 2005).

In the 1980s, the call for fairer evaluation practices further led researchers and advocates to rethink assessment tasks and conditions to make them more inclusive and support learners with special needs (e.g. dyslexia, dyspraxia). This research pushed the idea that appropriate accommodation could ultimately contribute to providing more equal educational opportunities for all (Kozleski & Yu, 2016).

These trends underscored the understanding that educational systems bear responsibilities to increase access, participation, and learning opportunities for whoever were the marginalised student populations at the time. They also marked a shift towards more sociocultural and sociopolitical views on assessment alongside the early psychometric views on evaluation (Elwood & Murphy, 2015).

Sociocultural and sociopolitical views on assessment argue that fairness is not merely a technical matter of refining variables, or ensuring that tests measure intended constructs. Instead, these approaches extend the notion of validity to include what Messick (1996) terms 'consequential validity' – the real-world effects that assessments have on individuals and society. From this standpoint, assessment practices are not neutral or universally applicable; they are shaped by cultural, historical and institutional contexts.

Elwood and Murphy (2015) further conceptualise assessment systems as 'cultural scripts' – socially and culturally embedded frameworks that guide how assessment is designed, interpreted, and enacted. These scripts reflect deeply rooted beliefs about knowledge, learning, and academic worth, which are passed down through generations and become normalised in educational practice. The role of sociocultural researchers, they argue, is to critically examine these inherited assumptions, question dominant assessment norms, and explore how such scripts may advantage or disadvantage different groups of students, raising important ethical and moral considerations.

Fairness in assessment and linguistic diversity

One dimension of cultural scripts that has become ingrained in European universities is that assessment design often sidelines the reality that students might have complex linguistic and cultural backgrounds. For generations, in fact, assessments were designed for the most part for students assumed to have monolingual, culturally homogeneous profiles rather than for diverse, plurilingual or mobile, cosmopolitan, international students (Shohamy, 2011). This does not mean that there were no conversations around linguistic diversity and inclusion, but the discussions tended to concern specific minorities (e.g. minority groups, second-generation migrants, etc.) in contexts where one language of instruction usually occupied hegemonic power. Researchers and policy-

makers paid less attention to the needs of plurilingual students who mastered the academic language of instruction to different degrees (Wang & Li, 2020). The research reflecting on the challenges of multilingual assessment in education (De Backer, 2020; Shohamy, 2011; Ysenbaert et al., 2020) and strategies to address these challenges is fairly recent (De Backer et al., 2017; Gorter & Cenoz, 2017; Schissel et al., 2018) and can be linked to increased student mobility and understanding of education as a global phenomenon.

As the university classroom has become more diverse, and education more internationalised (Knight, 2017), researchers have increasingly been calling attention to the fact that there are problems with using monolingual and monocultural norms and constructs to assess multilingual, diverse students (Menken & Shohamy, 2015; Shohamy, 2011).

These problems arise in part from misunderstandings about how learners acquire language and new understandings of what language is or does in a context of mobility and diversity of unprecedented scale (Budach & de Saint-Georges, 2017). As Blommaert and Backus (2013) argue, modern learners are most often plurilingual speakers, who have traversed many geographical, social and cultural contexts in their lifetime. As a result, many of them have not acquired a stable, native-like proficiency in more than one language in the traditional sense (Otsuji & Pennycook, 2010). Their complex linguistic repertoires can better be understood as composed of bites of languages and registers acquired while navigating these variegated contexts, and are often in various stages of development (Blommaert & Backus, 2013, p. 21). As a result, students' repertoires do not necessarily resemble the seemingly homogeneous native-like proficiency that the education system often expects (Weber & Horner, 2012).

This aspect of the multilingual experience poses a real challenge for traditional assessments. While learners' complex multilingual repertoires might be perfectly fitting to follow lessons, in assessments plurilingual learners might be more prone to misinterpret questions, be unfamiliar with vocabulary and cultural references, or struggle with adhering to valued standard forms of expression (De Backer, 2020). These challenges may, in turn, lead instructors to underestimate the abilities of plurilingual learners and result in unfair assessment.

Today, new calls for rethinking assessment (De Backer et al., 2017; Menken & Shohamy, 2015) and for developing policies and practices supportive of learners' multilingual competencies (Gorter & Cenoz, 2017) are emerging at all levels of education. In Higher Education, these calls appear alongside invitations to more substantially include the voice of students in reflections around curriculum design and university policies (Gerritsen-van Leeuwenkamp et al., 2018). The remainder of this paper, therefore, constitutes a deliberate effort to learn about students' lived experiences of assessment. We look at one Master's program at the University of Luxembourg as a prime context for the study, in order to bring to the surface some tacit assumptions about language and assessment in the international classroom and their unintended consequences for learners.

Methodology

Context

In the European landscape of Higher Education, the University of Luxembourg comes across as a hub for international policies and practices. Not only are more than 60% of

the University's students international students – a defining criterion of internationalisation according to Hawawini (2011) – but the university is also one of a handful of universities with an explicit multilingualism policy. While English serves as the main lingua franca of science for many disciplines, French, German, and Luxembourgish – the three official languages of the country – are also significantly represented in teaching, daily communication, research, and administration. Multilingualism and lingua franca use are thus common features of teaching and learning, with no single language holding hegemonic power.

Aligned with its multilingual policy, the university offers many bi- or multilingual study programs. One of them is the trilingual interdisciplinary masters in focus here, whose curriculum prepares Luxembourgish, EU, and non-EU students for careers as communication specialists in Luxembourg and the global market. The program is international, each year welcoming around 35 students from around the world – after a selection process that includes proof of meeting defined language criteria. Students must demonstrate (through diplomas, life experience, or international tests) a C1 level in English, a B2 level in one of the other program languages (French or German), and at least an A1 level in the third language (German or French). Besides their plurilingual profiles, students come with a broad range of academic experiences, mobility journeys and cultural backgrounds.

Instructors in the program tend to adopt a 'flexible multilingualism' orientation (Weber, 2014), viewing students' plurilingual repertoires as resources to build upon rather than dismiss. They may offer course materials in multiple languages, allow multilingual group discussions, teach bi- or multilingually, or permit students to write papers and assignments in the language they master best (e.g. writing a final essay in English for a course predominantly delivered in French).

Furthermore, students and instructors do not necessarily share the same varieties of English, French, or German. While instructors celebrate the diversity of linguistic practices in in-class conversations, standard varieties sometimes still remain the implicit and unquestioned norm in assignments (Canagarajah, 2006). This means that even instructors who theoretically recognise the benefits and reality of multilingualism find it hard to completely break away from a 'monolingual habitus' (Gogolin, 1997) in evaluation.

The diversity of the classroom requires instructors, therefore, to navigate two strong currents (Miller, 2017). On the one hand, instructors understand that students are often learners of the language in which they are instructed and 'should be given room to make mistakes as they engage in the difficult task of writing in a second (or sometimes third, fourth, or more) language'. On the other hand, instructors remain committed to upholding standards of clarity and correctness that are hallmarks of academic writing and that future employers will expect from students beyond their degrees, which can create tensions if instructors do not develop reflexivity on these normativities (de Saint-Georges et al., 2020).

Research design

The rationale for developing a research project on evaluation practices in this linguistic-cultural complex surges from my experience teaching in this program for the past 15

years and feeling an increasing need to understand the ways in which assessment interacted differentially within cohorts of such diverse profiles.

In the Spring of 2023, I decided to use a first-year introductory methodological course focused on 'research interviews' to seek out students' perspectives on the question of how instructors' perceived scripts affected, in their views, their outcomes – in order to understand better the ethical and moral impact of assessment design. For this class, I decided to design a joint research project in which we would all explore together the interaction between multilingualism and evaluation in our context of study, while learning how to conduct and analyse interviews in a real-life project, following a co-research design where students are viewed as partners (Clark et al., 2022).

For this project, a group of 27 students got involved. Students were tasked with interviewing a second-year peer from the program, following a common interview schedule we designed together. Each student sought informed consent from their interviewees and recorded and transcribed their exchanges. In the module, the students further learned to analyse interviews in groups, using a thematic analysis method (Braun & Clarke, 2006) before presenting their results to the class. In the final presentations, we critically discussed lessons learned about evaluation and the research process. In total, the students conducted and transcribed 23 semi-structured interviews, as a limited number of students chose the option of conducting their interviews in focus groups.

For this paper, I returned to the corpus built by the students and analysed the data in-depth myself. I used a ten-step systematic process relying on coding and various stages of discourse analysis to gain insights both within and across interviews (Dierckx de Casterlé et al., 2012). The data was thus constructed in multiple ways – as a piece of coursework which necessarily influenced the interviews, with a pre-set of questions in mind, and a complex intended audience (the students, the teacher and other stakeholders as the data was collected with permission to develop further research). Aligned with an interpretative and constructivist epistemology, the analysis thus does not suggest that students' narratives present 'objective' or 'truthful' views about evaluation in this multilingual context (Talja, 1999, p. 6). Rather, the analysis considers that the interview data opens a window into students' experiences, bringing to the surface the complexities of assessment in a multilingual context, as a starting point to initiate critical conversations and invite further research.

To focus the analysis, I also drew on the lense of assessment systems as cultural scripts, as outlined by Elwood and Murphy (2015). Because tacit assumptions passed down through generations are often difficult for those who design and present assessment tasks to recognise, I chose not to interview instructors, but instead to reconstruct the scripts from the perspective of the learners who are subject to them – scripts that influence their expectations and behaviours. To be sure, students' perception in the end might differ from what instructors had in mind. Nonetheless, their perspective open a window to discuss about the complexities of evaluation in multilingual contexts of education.

The analysis section considers first the challenges students express they face with evaluation and the operating scripts these challenges suggest in their view. Next, I turn to examine moments where students notice adjustments to the scripts and discuss whether they find them useful or not.

Findings

Scripts that challenge

As mentioned earlier, scripts can be defined as the social and cultural frames within which students and instructors operate. They show up in instructions as well as in how students' react to them. Across the interviews, the students mentioned difficulties that seemed to allude to three implicit scripts that could be identified in the classroom.

The monolingual proficiency script

A first script emerged in the data that we can call the monolingual proficiency script to refer to the view that even in multilingual contexts of education lingers the perspective that students share a stable, native-like proficiency in more than one language. The script can be summarised this way.

Script 1: Assessment can be designed without paying attention to language. Students master the language of instruction. They master spelling and grammar and are equipped with more or less the same linguistic resources to meet requirements.

Because the students are advanced master's students, and they have been admitted with certain language requirements to the program, the interview data suggests that language of instruction is considered as unproblematic whether to understand course content or to display what the students have learned or understood about this content, even in a context of great awareness of the plurilingualism of the students.

However, the students interviewed expressed struggling with language of instruction in different ways. Some students expressed struggling with the confident use of the language of instruction. Piotr³, for example, explained that while he can easily follow courses in the program, this ease does not mean he could always say 'what (he) wanted to say'. In contrast, Lisa does not have problems expressing herself in the language of instruction, but the challenge for her is the cost of this operation in terms of time. In French, aligning with standards takes her, for example, more time than in English:

- (1) I could probably write assignments in French with some grammar slash graphic mistakes, but French needs all the time. I need to go back, I need to check, dictionary, grammar, and it will take a longer time to prepare an assignment, so English is the easier one.

Other students expressed struggling with meeting standards, suggesting that the perceived standards the instructors set with regard to linguistic performance in the language of the course is challenging for students. Indeed, one important socialisation for students of language and communication is that spelling and grammar matter in academic writing. Yulia, for example, comes from a Bachelor's program where instructors typically evaluated students against a dominant language in the classroom/the country. In such a program, operating from a monolingual habitus (Gogolin, 1997) was the norm, and she explains it was common to 'count errors, linguistic errors', which would count towards the final grade. Because they are socialised into this script, students often chose to write in the 'safest' language. Steffi expresses the grade conundrum in this way:

- (1) Steffi: (...) I think in a way this master should encourage us to, to deal with all the languages we know and with our weakest languages in order to improve them. But as you said, sometimes then you have this grade up on your shoulder (...), waiting for you. (...) I think most of the people will always go for the easiest option.

Meeting standards is, to be sure, a goal that matters in an academic program centered on communication and language. Yet, some students pointed out how standards could also create disparities between students. As Steffi spells out, there can be an imbalance between students in the program:

- (2) Steffi: most of the time (...) most of us deal with languages that are maybe the third or even fourth language it's not taken that much into account. And so here comes the question of what is fair and what's not. (...) So if a German-speaking student is taking an assignment in German and, a student whose German is the third language (...) where's the balance?

The interviews showed, however, that sharing an L1 with the instructor does not guarantee successful outcomes, as Gail, re-telling the story of one of her classmates, reminds us that which standards are held also affects outcomes:

- (3) So, for example, one of my course mates was from [country in South East Asia]. She thought that she had a really high level of English, but in [country], like you know, they mix English with local languages. So it's kind of like let's say nonstandard variety of English. Or another one from [country in central Africa] had the same issue that she also, well, she's mainly a French-speaking person, but her French was not standard. So her papers were thrown back or they got bad grades because of that.

If we look across the interviews, students' views seem to corroborate what other research has shown for other contexts. For example, for primary school, De Backer (2020, p. 107) shows that for emergent bi- or multilingual students, doing evaluation tasks in a language less-mastered is often perceived to require a lot more time, attention, focus and energy, a point that might require some attention in designing evaluation, and which seems also relevant to higher education learners. Students also seem to perceive a more or less implicit linguistic market in the program – with some languages and varieties carrying more weight than others, and a monolingual habitus very much the benchmark at the point of evaluation. The excerpts suggest that in this complex, multilingual context of education, there are multiple layers and normativities at play which complicate outcomes for students as much as they are challenging for instructors.

The stable proficiency across modes script

A second script that circulates in the interviews concerns students' proficiency across modes of communication and can be formulated in the following way.

Script 2: Students who speak well (oral receptive and productive modes) also write well in the language of instruction (written production mode).

The students interviewed are enrolled in a trilingual course of study. While the program includes three official languages, students are not expected to be equally proficient in all of them. English is expected at the C1 level, German or French at B2, and the third language – either German or French again – at A1. Other languages are valued and may be used for group discussions or assignments, provided the instructor can evaluate work in that language – but in practice, their use remains limited, as they are not part of the program’s emphasised repertoire.

Regarding the official languages of the program, students are generally stronger in one or two of the languages of instruction and a few might have little mastery over the third one. They might also have what Blommaert and Backus (2013) call ‘partial competencies’ in any of these languages. For example, their reception (reading and listening) and production (talking, writing) skills might be unevenly distributed, meaning they might speak better than they write. Yet, the students seemed to identify an implicit script that because they participate a lot and express themselves well in the classroom, they will also encounter no problems in writing papers and essays. For Lorenz, however, thinking that students have evenly distributed abilities is a misleading assumption:

- (4) this friend (...) her English is very good? (...) But I could imagine that her written English is not as good as her spoken one (...). She had (...) a different exam, it was a paper. (...) And for her, it was a big disadvantage, because it was (...) graded on the academic English. Which is fine. You can say, Okay, I’m also going to grade your English. But some people are better, some people worse. So like, how can you in one assignment, change your academic English level? So it’s kind of unfair, because those people who are not good enough in English, they already [have] a worse grade

Students with strong writing skills in the language of instruction were not guaranteed a smooth journey, however. Some students expressed struggle not in writing, but in mastering the specific genre of academic discourse, as Francine highlights:

- (5) writing in the corporate world, and then changing into writing into the academic world, it’s a big challenge, a bit different (...) you have to adapt. Your practice, your brain and everything (...) to explain things in an academic way.

Francine later recalled how students not even trained in this way might meet even dire consequences:

- (6) some of my course mates had really, really hard times because they (...) never had to write a paper and they wrote words like ‘guys’ in academic paper, some of them actually failed courses because of that, and even though they knew the, whatever was required from them, they couldn’t write the papers. So, they didn’t qualify, and they failed, and they had to repeat.

Identifying whether students failed only because of a lack of mastery of the genre is a matter of empirical investigation. However, the multilingual context of evaluation raises questions such as: what advantages do those students whose repertoire overlaps with that of the program have? And do modalities of evaluation (oral, written) privilege

some students over others? Here again, the question is not simple and the response probably varies between cohorts, disciplines taught or learning outcomes envisioned.

The task types are familiar script

A third script that emerged in the data is that because students are masters' students, the tasks they need to do need no detailed explanations. They will be familiar with what instructors mean when they ask for certain types of assignments. This script can be summarised as:

Script 3: Students are familiar with evaluative tasks and instructors' expectations.

The data reveals, however, that many students interviewed feel insecure about the sorts of tasks/assignments requested. Thea, for example, explains how she was not familiar with the categories 'research paper' or 'personal essays' before entering the program. She came from an educational system where evaluation came solely in the form of oral exams, written tests, and rote memorising of content to be spilled back to the instructor. For her, 'to do critical writing and a reflexive, essays that was, difficult for me at the beginning'.

Lorenz felt some instructors, at times, could operate from an implied consensus. They gave instructions for assessment tasks as if everyone understood their expectations. They did not seem to take into account that in the internationalised classroom, students might not all share familiarity with the instructors' scripts of evaluation:

- (7) a friend of mine, she had a different (...) study background. So like, for example, writing articles, writing papers and stuff, either she never did, or she did in a different way. And, so, for her, it was really, really hard to meet the standards, because she just did not know how to do it, because nobody told her. Because of course, people expect in a Master, they expect that in the bachelor, you have to have this and this, knowledge and experience of writing papers, and it is not really explained (...) it is assumed that you kind of already know (...) but for the students, it's very frustrating and very challenging, because they just don't know how to. And then they feel like they're not good enough, because they are not smart enough, or they don't understand. But they just don't have the tools to do it.

Nancy further described the frustrating process of trying and sometimes failing to identify the expectations of the teacher in the absence of a shared script and detailed instructions, and how sometimes she would only understand the point of an assignment too late:

- (8) [in personal essays] the idea of, the professor I think is to receive different (...) visions of the subject. But it also makes everything more difficult because, everybody understands it in a different way and, you realise that only after you get your grades (...). Sometimes you are disappointed because. 'ah I understood it in a completely different way' and it was the wrong way, but before you couldn't realise it.

Lorenz also discussed an instance where the instructor dismissed students challenges with kind words of support, but provided little resource to meet assessment goals:

- (9) there was an assignment for which we didn't have much practice or much information. And for me, it was very easy, very easy. (...) I didn't have to prepare it. And I had 20 points (...) it was a very fair (laughs) and nice and good experience (laughs) of this exam. But for her, because there was no information, how to do it, what to do. (...) it was very unfair, because (...) there was no, practice leading up to this exam. It was just, you know, like, « you can you can make it », « you're gonna », you know, « you're gonna be fine. All of you, I know, you're gonna make it », you know, that was kind of the (...) (laughs) preparation, you know, like, make you guys feel like, you can do it. And then for her, it was not the case. So (...) I was really on her side. Because it is kind of unfair. If you don't have the tools to do it, and then someone just instead of preparing you says « yeah, yeah, you're going to be fine ».

Asking for instructions could be just another way for learners to try to please the instructors and get the 'recipe' handed to them, rather than putting in the hard work of learning. But research shows that 'assessment fog' (McCarthy, 2017) – not knowing what is expected – could also lead to misleading impressions about what the students know and can do because their texts are a shot in the dark. Other research has shown that when we expect learners to be autonomous and understand the rules by themselves, this often benefits those closest to the culture of the teacher (Reuter, 2007). While this challenge can also affect students whose L1 matches the language of instruction – and is not solely a matter of language or academic culture –, the interview data shows that is a particularly vivid concern among international students, who tend to experience it rather acutely.

Despite all these challenges, across the interviews, students also noticed many instances where instructors seemed to adjust their practices to support the students in meeting the intended outcomes. Below, I explore three such adjustments.

Adjustments that support

Recognising students profiles

Students in the interviews talked about how one form of adjustment came when instructors sought intentional information about students' experience with assignments. Piotr mentions an instructor, who instead of assuming the students mastered the language of instruction and the genre, explicitly asked them the question in the instruction for their final paper submission:

- (10) we had one course where the professor asked us to write if we had dealt with essays before and (...) what our proficiency in languages is. I think this is a good practice because a professor does not know what is your first foreign language. And if you say you have never written essays before or you write assignments in this language for the first time, it is a bit different situation from that when you write you are a professional journalist and prepare thousands of texts per year and it's your native language.

Nour noticed that some instructors offered the possibility to experiment with writing, allowing them to perform multilingually:

- (11) Okay, so I had one exam, where the teacher said we can write (...) in several languages. And I thought that was, really cool, because, in this exact exam, I was

kind of struggling to stay in one language. Because, when you hear stuff in English, and then when you analyse it, and I'm thinking in German, so it really helped me to, like, start the sentence in English and finish it in German, because it really was how I thought. So that was really helpful.

Isabelle, on the other hand, enjoyed the possibility offered by one instructor to write in languages not formally a part of the program, provided the students attached an automatic translation alongside the original paper. She really enjoyed that in her case she could also write in Spanish because the instructor was proficient in assessing papers in that language:

- (12) 'But [instructor's name], for example, I really liked the fact that she says 'I can read in English, French and Spanish, so if anyone wants to write in Spanish, they're welcome to do it'.

Students in the interview expressed that they appreciated when instructors showed care and valued their full repertoires. They considered that the process of evaluation was fairer when the instructor knew about their linguistic and cultural profile, even if they also recognised the challenge of factoring this information in grading papers.

Scaffolding the assessment journey

Another kind of adjustment students found beneficial, was when instructors did not leave them in an assessment fog but scaffolded the assessment process. They gave students cues to understand expectations and provided support to meet assessment goals. Aikerini relates also how receiving feedback supported successful outcomes in her view:

- (13) For the professors who gave me feedback last semester, I really appreciated it, and it helped me a lot to pay attention to those things that I maybe did wrong, and then I could improve it. (...) One other thing that I noticed is that the first works I submitted, they were not as good as the last ones, so in writing the essays, I improved myself, and then the first feedback I got, they helped me to work better [on the next one]

Other students found it valuable when instructors adopted an incremental approach to evaluation tasks. During his previous studies in the Middle East, Idriss was used to a system with only a final evaluation for a course 'and that's it'. Evaluation was always a high-stakes pass/fail moment. By contrast, he enjoys the fact that multiple tasks can build towards the final evaluation in his new program, 'where you get the final mark depending on all of those stages'. Such cumulative views allowed him to write, rewrite, edit the work and sharpen sentences and arguments.

Centering content over form

Finally, students writing in an additional language seemed to appreciate when instructors centered content and ideas over meeting stringent linguistic standards and 'put content into focus' (Amara), at least in the earlier semesters:

- (14) So, I think here, the professors, they are aware that the students are not native speakers. Most of them focus on your ideas, on your structure, and not on the linguistic

part. And it's (...) very helpful because some students, maybe they don't speak very well English, maybe their foreign languages are German or French.

Ida noted that she liked it when she had to do a presentation first (oral mode) before writing her essay (written mode), which she felt made her final production stronger. Alexandra further proposed adjusting the curriculum more generally. When asked what change she would propose for evaluation in the Master, she suggested that in the initial semesters 'she would start the master's with more basic introduction courses, and would be a little more flexible with the papers', focusing more on content than standards. In the more advanced semesters, she proposes more stringent rules: 'when you learn how to write academically in your language, then they can impose these academic points on your paper'. Another student, Francine, suggests a similar *modus operandi*:

- (15) So when you first come here in the first semester, [the] first half, of your studies, it would be better to have courses where you can (...) freely express yourself, and then gradually expose academic style on you (...) So it will be like changing courses from one semester to the other.

In the end, students considered that the processes of breaking assignments into parts and receiving feedback on each were useful in the international, multilingual classroom. They suggested that revealing the ropes of academic genres had a formative function and contributed to successful outcomes, and that scaffolding assessment processes helped meet diverse students' needs. For the students, this processual and developmental view of assessment allowed them to revise and refine their essays and papers over time, gradually aligning with the learning objectives and assessment criteria (Roy, 2016) – a feature that proves beneficial for all students, whether plurilingual or not.

Conclusion

This study set out to explore multilingual students' perceptions of assessment in an internationalised higher education context and to identify persistent assessment scripts shaping their experiences. In response to the first research question – what implicit assessment scripts shape students' experiences of evaluation – drawing on qualitative data from a co-research project in a trilingual master's program, the analysis revealed three pervasive scripts: (1) despite the multilingualism of the program, assessment tasks were perceived by students to be designed as if all students had stable, native-like proficiency in more than one language, treating the language of instruction as unproblematic for assessment, revealing the strong presence of a monolingual legacy even in a multilingual context of education (*monolingual proficiency script*); (2) despite the fact that students might have only partial competencies in the three languages of the program (e.g. speaking better than they would write academically), they felt some instructors still designed their task as if students would encounter no challenges with the different text types and genres of academic essays (*stable proficiency across modes script*), (3) despite the fact that students came to the program with varied background and educational trajectories, they felt assessment tasks were often still designed as if students shared familiarity with the academic task types they had to conduct (*task types are familiar*

script). The analysis revealed that these scripts often created tensions for students whose linguistic repertoires and educational backgrounds did not align with implicit expectations, leading to feelings of inadequacy, confusion and perceived unfairness.

Regarding the second question – do scripts create tensions or barriers to demonstrating knowledge – the students identified that even in a multilingual program, monolingual and monocultural assumptions persists, possibly disadvantaging students with complex linguistic profiles. The findings, however, also revealed that practices existed that contributed in students views to fairer assessment practices: instructors made adjustments that students perceived as supportive: eliciting information about students' linguistic backgrounds, scaffolding assignments, and prioritising content over form, at least in the earlier semesters of the programs.

These findings align with Elwood and Murphy's (2015) conceptualisation of assessment systems as cultural scripts, illustrating how when left unreflected, old standing beliefs and practices can inadvertently reproduce inequities. They also extend Shohamy's (2011) critique of monolingual assessment by providing empirical evidence of how multilingual repertoires interact with institutional norms. Moreover, the students' narratives reaffirm Blommaert and Backus's (2013) view of repertoires as fluid and context-dependent, challenging the validity of standardised expectations in diverse higher education classrooms.

Pedagogically, the study underscores the need for instructors to critically reflect on their own assessments scripts – where they originate, what values they encode, and how they align with the realities of increasingly diverse students populations. The linguistic composition of the classroom, more variegated blending multilingualism, *lingua francas*, and rich, complex cultural repertoires raise the question of how to make the classroom the best learning crossroads for multipolar views of the world, multiple language practices, and multiple languages to interact.

Ultimately, this research contributes to the growing body of work advocating for culturally and linguistically responsive assessment in higher education. It calls for further interdisciplinary inquiry into how multilingualism and internationalisation reshape the ethics and practice of evaluation, and how institutions can better support equitable learning outcomes in a global classroom, including by reflecting about normativities dominant in education (de Saint-Georges et al., 2020).

Reflecting together with students at the end of the course project, we recognised that maybe fairness was essentially an elusive goal. But, with Gipps and Stobart (2009), we agreed that – 'the best defense against inequitable assessment', is 'openness about design, constructs, scoring and grading', because reflecting on our scripts 'brings out into the open the values and biases of the [assessment] design process, offers an opportunity for debate about cultural and social influences, and opens up the relationship between assessor and learner' (p. 116). What was certain for us is that if universities want to do justice to the new constituency of their classrooms, they might do well to support more research that empirically investigates the tangled interactions between multilingualism and internationalisation in higher education, as well as cultivate new pedagogical practices, including for evaluation and assessment.

Notes

1. In 2022, around 6.9 million students were studying abroad, marking a 176 per cent increase over the past two decades (UNESCO Institute for Statistics (UIS, 2024)).

2. Following Moore and Gajo (2009), I use the term *multilingual* to refer to contexts where multiple languages co-exist, and *plurilingual* to refer to the individual repertoire of students who use several languages in their daily lives and studies.
3. Pseudonyms are used and details allowing to identify the students were removed.

Acknowledgements

I sincerely thank the editors and reviewers for their valuable feedback on an earlier version of the manuscript. The research project was approved by the Ethics Review Panel of the University of Luxembourg under the reference ERP 22-079.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- AERA. (2014). *Standards for educational and psychological testing*. American Educational Research Association. ISBN 978-0-935302-35-6.
- Anderson, T. R., & Rogan, J. M. (2010). Bridging the educational research-teaching practice gap. *Biochemistry and Molecular Biology Education*, 38(1), 51–57. <https://doi.org/10.1002/bmb.20362>
- Blommaert, J., & Backus, A. (2013). Superdiverse repertoires and the individual. In I. de Saint-Georges, & J. J. Weber (Eds.), *Multilingualism and multimodality* (pp. 11–32). Springer.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Budach, Gabriele, & de Saint-Georges, Ingrid (2017). Superdiversity and language. In Suresh Canagarajah (Ed.), *The Routledge Handbook of Migration and Language* (1st Edition, pp. 63–78). Routledge.
- Canagarajah, S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly*, 3(3), 229–242. https://doi.org/10.1207/s15434311laq0303_1
- Clark, A. T., Ahmed, I., Metzger, S., Walker, E., & Wylie, R. (2022). Moving from co-design to co-research: Engaging youth participation in guided qualitative inquiry. *International Journal of Qualitative Methods*, 21. <https://doi.org/10.1177/16094069221084793>
- De Backer, F. (2020). Bridging the gap between learning and evaluation. *Journal of Applied Linguistics and Professional Practice*, 14(1), 96–118. <https://doi.org/10.1558/jalpp.39770>
- De Backer, F., Van Avermaet, P., & Slembrouck, S. (2017). Schools as laboratories for exploring multilingual assessment policies and practices. *Language and Education*, 31(3), 217–230. <https://doi.org/10.1080/09500782.2016.1261896>
- de Saint-Georges, I., Budach, G., & Tress, C. (2020). “We need to become ‘educational chameleons’”: From unified to multiple norms in a multilingual and international higher education context. *European Journal of Applied Linguistics*, 8(2), 233–256. <https://doi.org/10.1515/eujal-2020-0011>
- Dierckx de Casterlé, B., Gastmans, C., Bryon, E., & Denier, Y. (2012). QUAGOL: A guide for qualitative data analysis. *International Journal of Nursing Studies*, 49(3), 360–371. <https://doi.org/10.1016/j.ijnurstu.2011.09.012>
- Elwood, J., & Murphy, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182–192. <https://doi.org/10.1080/0969594X.2015.1021568>
- Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2018). Developing questionnaires to measure students’ expectations and perceptions of assessment quality. *Cogent Education*, 5(1), 1–16. <https://doi.org/10.1080/2331186X.2018.1464425>

- Gipps, C., & Stobart, G. (2009). Fairness in assessment. In C. Wyatt-Smith, & J. J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 105–118). Springer. https://doi.org/10.1007/978-1-4020-9964-9_6
- Gogolin, I. (1997). The “monolingual habitus” as the common feature in teaching in the language of the majority in different countries. *Per Linguam*, 13(2), 38–49. <https://doi.org/10.5785/13-2-187>
- Gorter, D., & Cenoz, J. (2017). Language education policy and multilingual assessment. *Language and Education*, 31(3), 231–248. <https://doi.org/10.1080/09500782.2016.1261892>
- Hawawini, G. (2011). The internationalization of higher education institutions: A critical review and a radical proposal. *INSEAD Working Paper (2011) 112/FIN*. <https://doi.org/10.2139/ssrn.19Hu54697>
- Hu, A. (2018). Les universités comme espaces de communication interculturels et multilingues: Pourquoi le discours sur l’internationalisation devrait être plus fortement imbriqué avec le thème du multilinguisme. In J. Erfurt, A. Weirich, & E. Caporal-Ebersold (Eds.), *Education plurilingue et pratiques langagières. Hommage à Christine Hélot* (pp. 95–116). Peter Lang.
- Knight, J. (2017). The international university: Models and muddles. In R. Barnett, & M. A. Peters (Eds.), *The idea of the university (Vol.2): Contemporary Perspectives* (pp. 99–119). Peter Lang.
- Kozleski, E. B., & Yu, I. (2016). Inclusive education. In L. Meyer (Ed.), *Oxford bibliographies in education*. Oxford University Press. <http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0162.xml>
- Lam, T. C. M. (1995). Fairness in performance assessment: ERIC digest [Online]. <http://ericae.net/db/edo/ED391982.htm> (ERIC Document Reproduction Services No. ED 391 982)
- McCarthy, J. (2017). *So all can learn: A practical guide to differentiation*. Rowman and Littlefield. 978-1-4758-2572-5.
- Menken, K., & Shohamy, E. (2015). Invited colloquium on negotiating the complexities of multilingual assessment. *Language Teaching*, 48(3), 421–425. <https://doi.org/10.1017/S0261444815000166>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>
- Miller, E. (2017, August 24). Evaluating and grading multilingual writing. Supporting faculty, instructional staff, and teaching assistants as they teach with writing. <https://dept.writing.wisc.edu/wac/evaluating-and-grading-multilingual-writing/>
- Moore, D., & Gajo, L. (2009). Introduction – French voices on plurilingualism and pluriculturalism: Theory, significance and perspectives. *International Journal of Multilingualism*, 6(2), 137–153. <https://doi.org/10.1080/14790710902846707>
- Otsuji, E., & Pennycook, A. (2010). Metrolingualism: Fixity, fluidity and language in flux. *International Journal of Multilingualism*, 7(3), 240–254. <https://doi.org/10.1080/14790710903414331>
- Reuter, Y. (Ed.). (2007). *Une école Freinet: Fonctionnements et effets d’une pédagogie alternative en milieu populaire*. L’Harmattan. 978-2-296-03315-3.
- Roy, S. (2016). Assessing bilingual and multilingual learners in mainstream classrooms. In S. Scott, D. E. Scott, & C. F. Webber (Eds.), *Leadership of assessment, inclusion, and learning* (pp. 225–241). Springer. https://doi.org/10.1007/978-3-319-23347-5_9
- Schissel, J. L., Leung, C., López-Gopar, M., & Davis, J. R. (2018). Multilingual learners in language assessment: Assessment design for linguistically diverse communities. *Language and Education*, 32(2), 167–182. <https://doi.org/10.1080/09500782.2018.1429463>
- Secada, W. G. (1989). Educational equity versus equality of education: An alternative conception. In W. G. Secada (Ed.), *Equity and education* (pp. 68–88). Falmer Press.
- Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *The Modern Language Journal*, 95(3), 418–429. <https://doi.org/10.1111/j.1540-4781.2011.01210.x>
- Spolsky, B. (2008). Language assessment in historical and future perspective. In N. H. Hornberger (Ed.), *Encyclopedia of language and education* (pp. 2570–2579). Springer. https://doi.org/10.1007/978-0-387-30424-3_192
- Stobart, G. (2005). Fairness in multicultural assessment systems. *Assessment in Education: Principles, Policy & Practice*, 12(3), 275–287. <https://doi.org/10.1080/09695940500337249>
- Talja, Sanna (1999). Analyzing qualitative interview data. *Library & Information Science Research*, 21(4), 459–477. [https://doi.org/10.1016/S0740-8188\(99\)00024-9](https://doi.org/10.1016/S0740-8188(99)00024-9)

- Tierney, R. D. (2016). Fairness in educational assessment. In M. A. Peters (Ed.), *Encyclopedia of educational philosophy and theory* (pp. 1–6). Springer. https://doi.org/10.1007/978-981-287-532-7_400-1
- Wang, Y., & Li, S. (2020). Issues, challenges, and future directions for multilingual assessment. *Journal of Language Teaching and Research*, 11(6), 914–919. <https://doi.org/10.17507/jltr.1106.06>
- Weber, J. J. (2014). *Flexible multilingual education: Putting children's needs first*. Multilingual Matters. ISBN 9781783091980.
- Weber, J.-J., & Horner, K. (2012). *Introducing multilingualism: A social approach*. Taylor & Francis. ISBN 9781138244498.
- Ysenbaert, J., Houtte, M. V., & Avermaet, P. V. (2020). Assessment policies and practices in contexts of diversity: Unravelling the tensions. *Educational Assessment, Evaluation and Accountability*, 32(2), 107–126. <https://doi.org/10.1007/s11092-020-09319-7>